

the association for computational heresy

presents

a record of the proceedings of

SIGBOVIK 2017

*the eleventh annual intercalary robot dance party in celebration
of workshop on symposium about 2⁶th birthdays; in particular,
that of harry q. bovik*

carnegie mellon university

pittsburgh, pa

april 0, 2017



Association for Computational Heresy

Advancing computing as Tomfoolery & Distraction

SIGBOVIK

A Record of the Proceedings of SIGBOVIK 2017

ISSN 2155-0166

April 0, 2017

Copyright is maintained by the individual authors, though obviously this all gets posted to the Internet and stuff, because it's 2017.

Permission to make digital or hard copies of portions of this work for personal use is granted; permission to make digital or hard copies of portions of this work for classroom use is also granted, but seems ill-advised. Abstracting with credit is permitted; abstracting with credit cards seems difficult.

Additional copies of this work may be ordered from Lulu; refer to <http://sigbovik.org> for details.



SIGBOVIK 2017

Message from the Organizing Committee

These are the proceedings of the 1011th annual conference of the Special Interest Group on Harry Q.¹ Bovik, organised by the Association for Computational Heresy in honour of Harry Q. Bovik's 2^{6.20945...}th birthday.

2017 was a monumental year for SIGBOVIK. We experienced a factor of 1,522 increase in submissions over last year, that's to say, we received 35,000 submissions. For your convenience, we made a histogram showing the number of submissions across time in Figure 1. As much as we would have liked to accept all of these scientific breakthroughs, they would have caused our proceedings to weigh about 868 lbs or 394 kg,² and face it, nobody needs that heavy of a paper paperweight.

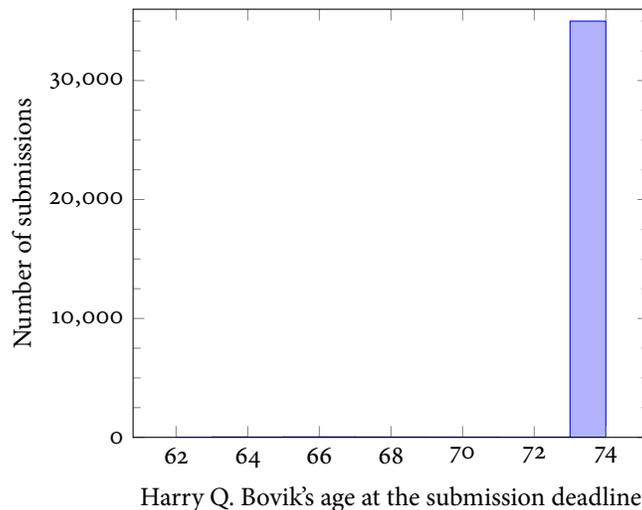


Figure 1: Number of SIGBOVIK submissions across time

Our reviewers tirelessly reviewed these submissions, and concluded that only ϵ of them were truly worthy of being accepted for inclusion in the SIGBOVIK proceedings. “What, if anything, is epsilon?”, you might ask. In this case, it so happens that $\epsilon = 0.001$, a value of ϵ found in the wild in the C++ species of programs 5.4% of the time, according to seminal work published in the proceedings of SIGBOVIK

¹In case you were wondering: Quahaug.



²

Assuming a page weight of 4.5 grams, double-sided printing, and an average of 5 pages per submission. This corresponds to a bit more than the weight of a Kodiak bear (*Ursus arctos middendorffi*) [Wik17], pictured left.

2014 [PhD14, Figure 6]. Comparing the SIGBOVIK 2017 acceptance rate with the acceptance rates of other “prestigious” computer science conferences (Table 1), we see that SIGBOVIK is anywhere from 169 times (SIGCOMM) to 283 times (SODA) more prestigious than other top computer science conferences.

Conference		
Name	Year	Acceptance rate
FOCS	2016	27.7%
ICML	2016	25%
POPL	2016	23%
SODA	2016	28.3%
SOSP	2015	17%
STOC	2016	24.9%
SIGCOMM	2016	16.9%
VLDB	2016	21.2%

Table 1: Acceptance rates at “prestigious” computer science conferences

SIGBOVIK used a new submissions website this year. Woven from the finest spider webs by Jordan Brown and Jean Yang, this website survived a DoS attack from the PC^W^W^W^W^W^W^W^W^W successfully handled 35,000 totally genuine submissions. We are grateful for their support and for letting us stress-test their software.

We hope that you will find the seminal works below informative and illustrative of the high-calibre research SIGBOVIK has become known for. From breakthroughs in debugging to new advances in impure math and game theory, we are sure there will be something of interest for everybody.³

Our thanks to the volunteers who made SIGBOVIK possible. In particular, we would like to thank Sol Boucher for assembling these proceedings; Rose Bohrer, Stefan Muller, and Ben Blum for reviewing papers and ensuring SIGBOVIK accepts only the highest calibre research; Carlo Angiuli for maintaining the SIGBOVIK website and his helpful advice; Chris Yu for the artwork; Catherine Copetas for managing SIGBOVIK finances and other administrative concerns; Ryan Kavanagh for organising the organisers; and last, but not least, the authors, without whom none of this would be possible.

The SIGBOVIK 2017 Organising Committee
Pittsburgh, PA

References

- [PhD14] Dr. Tom Murphy VII Ph.D. “What, if anything, is epsilon?” In: *A Record of the Proceedings of SIGBOVIK 2014*. Apr. 2014, pp. 93–97.
- [Wik17] Wikipedia. *Kodiak bear*. Mar. 2017. URL: https://en.wikipedia.org/wiki/Kodiak_bear.

³And if you can’t find something you’re interested in, you should have submitted it!

These papers are so awesome few can bear to look away!

Bear track: Strong Accept	3
1 Is this the shortest SIGBOVIK paper?	4
2 I'm on vacation so I'm submitting a vacation picture instead of a paper, or, perhaps, a vacation photo in the format of a paper; I hope a predatory open access journal e-mails me about this article	6
Mule track: I'm Going to Use This!	7
3 Who sorts the sorters?	8
4 Objectionability: A computational view of mathematical computation . . .	12
5 Towards a well-defined and secure flirtation protocol	15
6 A solution to the two-body problem	22
Magpie track: Who Said Money Can't Buy...	27
7 Call For Partners: Romance with rigor	28
8 Nano electronic transplant in brain and eyes to analyze and process human thoughts and emotions	34
9 Grant proposal: Monetary policy of sparkly things	36
Monkey track: It's Only a Game Theory	39
10 Is it Percival time yet?: A preliminary analysis of Avalon gameplay and strategy	40
11 Dr. Boozehead, or How I learned to stop worrying and get drunk: Design principles and analysis of drinking games in the silicon age	47
12 A boring follow-up paper to "Which ITG stepcharts are turniest?" titled, "Which ITG stepcharts are crossoveriest and/or footswitchiest?"	54
Dog track: Nonstandard ML	63
13 Batch normalization for improved DNN performance, my ass	64
14 Colonel density estimation	66
15 Degenerative adversarial networks	68
16 Stopping GAN violence: Generative unadversarial networks	76

Groundhog track: Putting the “Under” in “Image Understanding”	83
17 DeepDoggo: Learning the answer to “Who’s a good dog?”	84
18 □□□□□□□□□□□□□□ □□□□□□ □□□□ □□□□□ □□□□□□ □□ □□□□□□	88
19 Distinguishing humans from other forms of cattle	94
Chipmunk track: New and “Improved” Languages	101
20 On the Turing completeness of MS PowerPoint	102
21 Effective multi-threading in Befunge	107
22 Automatic distributed execution of LLVM code using SQL JIT compilation	114
23 WysiScript: Programming via direct syntax highlighting	119
24 ZM~ # PRinty# C with ABC!	129
Insect track: Debugging	149
25 Amazon Web Services: Field observations related to arachnid cohabitation .	150
26 Blackberry Debugging	153
Moose track: Impure Math and \Big Data	157
27 Fake news logic	158
28 RRR for UUU: Exact analysis of pee queue systems with perfect urinal etiquette	163
29 The next 700 type systems	169
30 A modular approach to state-of-the-art big data visualization	172
31 Efficient computation of an optimal portmantout	176
Raccoon track: Talkin’ Trash	191
32 Garbage collection for heaps only a mother could love	192
33 A new paradigm for robotic dust collection: Theorems, user studies, and a field study	194
Cat track: Work-stealingsaving	199
34 The Zero-color Theorem: An optimal poster design algorithm	200
35 Cerebral genus: Dead duck or phoenix?	203



Bear track

Strong Accept



- 1 **Is this the shortest SIGBOVIK paper?**

Joe Doyle

- 2 **I'm on vacation so I'm submitting a vacation picture instead of a paper, or, perhaps, a vacation photo in the format of a paper; I hope a predatory open access journal e-mails me about this article**

Jim McCann

Is This the Shortest SIGBOVIK Paper?

Joe Doyle

March 7, 2017

Maybe not.



CONFIDENTIAL COMMITTEE MATERIALS

SIGBOVIK 2017 Paper Review

Paper 14: Is This the Shortest SIGBOVIK Paper?

Reviewer ϕ

Rating: e/π

Confidence: ϵ

While the importance of this work to the SIGBOVIK community can't be debated (because we have no established protocol for such a debate), it is not novel. The content of this paper is entirely contained within the papers "The Portmantout" from SIGBOVIK 2015, "A shortmantout" from SIGBOVIK 2016 and "Efficient Computation of an Optimal Portmantout" from the present SIGBOVIK. It should therefore only be accepted if our goal is to increase our page count which, as always, it is.

I'm on Vacation So I'm Submitting A Vacation Picture Instead Of A Paper, Or, Perhaps, A Vacation Photo In The Format Of A Paper; I Hope A Predatory Open Access Journal E-Mails Me About This Article.

Jim McCann*
TCHOW llc



Figure 1: *At the “Inventions” buffet in Hotel Disneyland.*

*e-mail: ix@tchow.com



Mule track

I'm Going to Use This!



3 Who sorts the sorters?

Alexander R. Frieder

4 Objectionability: A computational view of mathematical computation

Cameron Wong and Dez Reed

5 Towards a well-defined and secure flirtation protocol

Rowan Copley

6 A solution to the two-body problem

Rose Bohrer

Who sorts the sorters?

Alexander R. Frieder*

March 31, 2017

1 Introduction

Any introductory programming class will cover the basics of different sorting algorithms. Sorting algorithms are the perfect introduction to algorithmic analysis for a multitude of reasons. First, there are straightforward proofs of the lower-bound in terms of performance, thus we can easily teach students that $O(n \log(n))$ sorts are optimal algorithmically. Second, the naïve sorts are fairly far from optimal and that is simple to show students. Finally, sorting is a deceptively easy task for humans. To expand on that final point, given a list of numbers, it is very easy for a human to order them correctly. Indeed, most introductory lessons will have the teacher prompting the student to describe exactly *how* they put them in order.

The important takeaway from the lesson is that sorting is not a trivial process and it must be done through some algorithm. **It does not suffice to take numbers and simply *put them in order*.** The lesson will then proceed by walking the student through multiple algorithms and then displaying their performance on some large set of numbers, presenting how long each algorithm took. The teacher will then often conclude with some statement like “Thus we can see that merge sort is faster than selection sort”. However, this directly contradicts the previous lesson: the teacher has taken performance numbers and somehow magically put them in order. Indeed, we need to use a sorting algorithm to determine which sorting algorithm is best. But which sorting algorithm is best for that task?

In this paper, we will review a variety of sorting algorithms and evaluate their performance on large randomly sorted arrays of numbers. We will then use the more general version of sorting to compare each sorting algorithm with each other to correctly evaluate which sorting algorithm is the best for sorting sorting algorithms. Finally, we will use the same algorithms to determine which sorting algorithm is best for determining which sorting algorithm is best for determining which sorting algorithm is best.

2 Sorting Algorithms

In this section we will review six different sorting algorithms that will be used throughout this paper and present, without proof¹, their algorithmic complexity. If you feel you are an expert sorter, feel free to proceed to section 3. I will not be offended. Not at all. I promise. These are not tears. I am just cutting onions.

2.1 Selection Sort

Selection sort is often the first sorting algorithm taught and is often the way humans intuitively sort. The algorithm behind selection sort is absurdly simple: find the smallest element left in the list, remove it, and put it in the output list. Repeat. Once the original list is empty, the output list will be sorted.

Since selection sort requires finding the max at each step, it runs in $O(n^2)$ time.

*email: alex.frieder@gmail.com

¹Unless you count Wikipedia as proof.

Exercise 1

Teach your dog² how to perform selection sort.

2.2 Insertion Sort

Insertion sort is a similar idea to selection sort, but instead of putting in the effort to pull the correct element from the original list, we put in the effort to put a random element from the original list into the right location. For each element in the original list, we remove it, iterate through the output list, find the smallest element larger than it, and insert it immediately before that. Thus, at each step, the output list is sorted, and at the end, the original list is empty and the output list is sorted.

Since inserting an element may require iterating over the entire output list, insertion sort runs in $O(n^2)$ time.

Exercise 2

Put the words of this paper into alphabetical order using insertion sort³.

2.3 Odd-Even Sort

Odd-even sort is a parallelizable version of the notorious and often politicized bubble sort⁴. In odd-even sort, we first iterate through the odd indices and for each element that is larger than the next element, we swap them. We then repeat for all the even indices. Finally, we repeat this pair of steps until no swaps are made. By definition, the list is now sorted. Since each step only swaps either exclusively odd or exclusively even indexed elements with its neighbor, the entire step can be completely parallelized.

Despite being parallelizable, odd-even sort is not more efficient than bubble sort and runs in a worst case of $O(n^2)$.

²Other pets may suffice for this exercise, but not cats. Cats are fundamentally incapable of learning algorithms.

³Ensure you have at least 15 minutes of free time before attempting this.

⁴https://www.youtube.com/watch?v=k4RRi_ntQc8

Exercise 3

There actually exists exactly one list of integers on which odd-even sort does not terminate. Find that list, treat the numbers as binary, read off the ASCII-encoded message, and continue your quest to find the Fountain of Youth.

2.4 Comb Sort

Comb sort is also a generalization of the notorious and often politicized bubble sort⁵. Comb sort is parametrized by a so-called shrink factor k , optimally around 1.3⁶. Comb sort begins by iterating over every pair of elements n away from each other, where n is the length of the list, and swapping the elements if they are not sorted relative to each other. Initially, that is only the first and last element. The distance is then set to n/k and the process is repeated. This process repeats, dividing the distance by k until the distance is 1, at which point we revert to bubble sort, swapping out of order pairs until the list is sorted.

Despite having faster run time than bubble sort, comb sort matches its complexity of $O(n^2)$.

Exercise 4

Do 30 push-ups, 30 sit-ups, and wall sit for at least 2 minutes.

2.5 Merge Sort

Merge Sort is the prototypical divide and conquer algorithm. First, recursively sort the first half of the list and the second half of the list, then merge them, by removing the smaller head of the two lists and inserting it into an output list. Repeat until both lists are empty. The output list is now sorted.

Merge sort takes exactly $\Theta(n \log(n))$ comparisons and is thus algorithmically optimal.

⁵https://youtu.be/M0zg_Cf4K4w?t=21s

⁶As cited from the Epic of Gilgamesh.

Exercise 5

Write an angry letter to the author explaining that the $O(n \log(n))$ lower bound only applies to comparison-based sorts and that some types can be sorted into linear time so therefore the author is wrong.

2.6 Quicksort

Quicksort takes, in the worst case, $O(n^2)$, but on average, $O(n \log(n))$.

Exercise 6

Enroll at Carnegie Mellon and take any of the following courses to learn in-depth about quicksort:

- 15-122
- 15-210
- 15-359
- 15-451

3 Methods/Implementation

For comparing sorting algorithms, the naïve method is to sort a large number of random lists of numbers and take the average time spent sorting. As mentioned originally, this method is antithetical to the entire purpose of sorting algorithms. But it will suffice as a first level approximation of the efficiency of sorting algorithms.

When comparing higher-order sorting algorithms we need to enforce a partial ordering. In other words, we need to define a $<$ operator on algorithms. There are multiple ways to do so, but we have chosen to use the most natural ordering: when comparing two algorithms A and B , we define $A < B$ to be true iff the total time spent sorting n random shuffles of some list L is less when sorting under A than when sorting under B . This is obviously a probabilistic total ordering, but in the limit, it is a deterministic total ordering. We will pretend our numbers are big enough to be considered “the limit”.

We used $n = 3$ across all experiments. For sorting algorithms, we used $L = [1, 10^5)$ and defined $<$ as the normal integer $<$. For sorting sorting algorithms, we used $L =$ the list of all sorting algorithms discussed in section 2. As described above, we determined for two algorithms A and B if $A < B$ by running A on 3 random shuffles of $[1, 10^5)$ and B on 3 random shuffles of $[1, 10^5)$ and using normal integer $<$ on the time taken. Finally, for sorting sorting sorting algorithms, we use the same L , that is, all algorithms discussed above. We define $<$ in the logical expansion of the previous definition: $A < B$ if 3 runs of A sorting sorting algorithms takes less time than 3 runs of B sorting sorting algorithms.

These were all implemented in Python 3.6 but ported to and eventually run in PyPy2.7 v5.6.0 due to time concerns.

4 Discussion

All of the numeric results of these tests can be found in table 1 on the next page.

For the integer sorting algorithms, we see exactly the results usually presented in introductory classes and as expected. Quicksort, using probabilistic magic, is fastest, with merge sort following closely. Comb sort, using its heuristic advantage, is fairly fast, with the other $O(n^2)$ algorithms following behind. There is nothing too exciting here. The full ordering is quicksort $<$ merge sort $<$ comb sort $<$ insertion sort $<$ selection sort $<$ odd-even sort.

However, both the sorting sorting algorithms and the sorting sorting sorting algorithms have a different order from the integer sorting: merge sort $<$ insertion sort $<$ quicksort $<$ selection sort $<$ odd-even sort $<$ comb sort.

There are two interesting changes here. We can use merge sort as a good baseline since it does a relatively deterministic and static number of comparisons.

First, insertion sort is much much better for higher-order sorting than for integer sorting. For integer sorting, insertion sort is about 60x slower than merge sort. However, for sorting sorting,

Algorithm	0th Order Time	1st Order Time	2nd Order Time
Quicksort	0.06127 s	329.88 s	18.62 min
Merge Sort	0.06757 s	277.32 s	16.65 min
Comb Sort	0.09867 s	792.78 s	235.46 min
Insertion Sort	4.03476 s	314.31 s	17.67 min
Selection Sort	8.15739 s	565.02 s	25.33 min
Odd-Even Sort	15.79098 s	781.93 s	216.02 min

Table 1: The time take for algorithms to sort integers (0th order), sorting algorithms (1st order), and sorting sorting algorithms (2nd order)

it is only 1.13x slower, and for sorting sorting sorting, it is only 1.06x slower. It is conceivable that since insertion sort spends most of its time iterating over the beginning of the sorted list that when comparisons get more expensive, it gets more efficient.

Second, comb sort is much much slower for higher-order sorting than for integer sorting. For integer sorting, comb sort is only about 1.5x slower than merge sort, but for sorting sorting, it is about 3x slower, and for sorting sorting sorting, it is 14x slower. It is the slowest sort for higher-order sorting and took almost 4 hours compared to merge sort’s 17 minutes. It is conceivable that comb sort, with its decreasing window, spends more time doing redundant tests as it reduces to bubble sort.

We ran tests counting comparisons to test both of these hypotheses and the evidence is clear: there is no evidence for these hypotheses whatsoever. Thus we present no explanation and offer a 50¢ prize for the first plausible explanation submitted to the author.

5 Conclusions

Use quicksort, unless you need to sort algorithms, in which case use merge sort. Most sorts behave the same across orders, but some do not. Thinking about higher-order sorting algorithms may cause temporary⁷ insanity.

⁷At least, I hope it is temporary...

6 Future Work

There are numerous ways this novel direction of research can be expanded:

- Higher-order functions including, but not limited to, 3rd, 4th, and ω th order sorting algorithms
- Something involving quantum computers working on big data in the cloud
- Determining if there exist sorting algorithms that excel at non-constant comparison algorithms⁸
- Other meta-analyses such as:
 - A search algorithm for searching for search algorithms
 - A database for managing databases
 - A container for containing other containers⁹
 - A cache invalidator for managing cache invalidation algorithms

7 Acknowledgements

The author would like to thank a handful of people, including Arley Schenker and Ben Lichtman for proofreading this and sharing in the madness, Thomas Bayes, for obvious reasons, the Academy, for making this all possible, and readers like you, without whom this whole ordeal would have been meaningless.

⁸This is actually an interesting problem that I could not find much research on.

⁹This may be made redundant by Docker.

Objectionability: A Computational View of Mathematical Computation

Cameron Wong, Dez Reed

March 10, 2017

Abstract

Any student in an “engineering” discipline (that is, the fields of math, computer science, mechanical engineering, underwater basket weaving, et al) can sympathize with looking at an equation or problem and realizing that pages of unintelligible scribbles would be required to reach a satisfactory conclusion. Similarly, professors often add allowances for “algebraic grunt work” that is often assumed to be correct in grading rubrics. However, such terms are often, at most, vacuously defined and it is up to the subjective preferences of whoever is looking to determine whether four pages of equational reasoning is sufficient or to remove points for omitting a further two pages of integration by parts, second-order substitution and a re-derivation of the Cauchy-Schwarz inequality.

In this paper, we attempt to provide a mathematical model to describe the “grossness” of any particular mathematical problem P (the precise formal definition of which is semi-intentionally left vacuous). In particular, we define $\mathbb{G}(P)$ to be the *computational grossness* of P and prove several further useless results that come of direct consequence. As such, we also establish several *grossness classes* for problems.

Finally, we introduce the concept of *objectionability*, whether any *arbitrarily gross* problem is uncomputable by a non-idealized human. Among other things, we will find that O , determining whether a problem is objectionable, is itself objectionable.

1 Introduction

In 1936, Alan Turing introduced the concept of a *Turing Machine* that could be used to compute formally-defined mathematical functions. In the same vein, the idea of *decidability* was introduced, determining whether a given problem was solvable with pen-and-paper mathematical manipulations.

We see, though, that some problems, while perhaps decidable, are simply “gross”- that is to say, they are difficult to generate solutions to. This phenomenon has even been referenced by the famed Jay-Z, referring to “99 problems” that seemed to be of trivial value compared to the problems that the song’s subject had been presented with.

However, in reasoning about decidability, little to no thought is given to whether anyone would *want* to solve such problems using pen-and-paper mathematical methods – with the advent of computers, many of these “gross” computations can be done with little to no effort on the part of a human. Even so, the fact remains that such “gross” problems exist. Further thought reveals that (without loss of generality) the problem of writing Python code to solve something may be “too gross”.

This paper attempts to resolve this issue by defining $\mathbb{G}(P)$, the grossness of P for a problem P in \mathbb{E} (the set of all problems) then discuss several schemes by which \mathbb{G} may be estimated. We also briefly discuss how we might determine whether a problem is *objectionable*, that someone may reasonably object to being asked to solve it.

2 Repugnance Theory

We begin by defining $\mathbb{G}(P)$ for some $P \in \mathbb{E}$ as the minimum grossness for any solution to P . As we will see, however, this is difficult to determine in a general case.

Consider the following problem:

Problem 1. Fix some alphabet Σ and other arbitrary set Π . Define the following sets:

$$\Gamma = \{(\mu, n, \tau) \in \{w \in \Sigma^* : w \text{ every proper prefix of } w \text{ is in } \mu\} \times \mathbb{N} \times \{\zeta \in \Pi : \langle \zeta \rangle \geq n\} : |\langle \tau \rangle|^{\mu} \langle n \rangle| \leq |\mu|^{\langle n \rangle} + |\langle \tau \rangle|\}$$

$$\kappa = \{(\Sigma^*, \iota, \psi) : \psi \in \{f : \Sigma^* \times \Sigma^* \rightarrow \Sigma^* \mid \\ \forall x, y \in \Sigma^*, f(x, y) = f(y, x)\} \text{ and} \\ \forall x \in \Sigma^*, \psi(\iota, x) = x = \psi(x, \iota)\}$$

where $\langle x \rangle$ is an arbitrary encoding of an object x over Σ^* .

Let Q be the subproblem of “Given a string S over Σ^* , is S equivalent to $\langle(\nu, \chi, \omega)\rangle$ where $(\nu, \chi, \omega) \in \Gamma \cap \kappa$?”.

Is Q decidable?

A naive approach might have a student attempt to determine what each set *means* and to reason about the properties and behavior of each. Note, however, that $\langle P \rangle$ over the alphabet of unicode characters occurring in this paper (which we will take to be the problem statement as given) contains 39 characters when any symbols that are not considered “common English characters” are removed. This, too, is a vacuous definition, however, so we will define a “common English character” to be anything that has a standard unicode codepoint less than 95. With this interpretation, then, we could say that $\mathbb{G}(P) = 39$, as any solution to this problem must begin by reading the problem statement. We will call this approach the “character counting” approach.

An alternative approach, however, might involve inspecting the types of the sets involved – μ is a string, n is a natural number and τ takes some type t , where t is the type of “things in Π ”. This is not to start a discussion of type theory, so we suffice to say that this approach leads to finding that $\mathbb{G}(P) = T$, where T is the set of all types that fulfill the given criteria. It then follows that the “character counting” method is only sufficient to provide an upper bound on \mathbb{G} .

A real type-chasing approach, however, would note that the definition of κ is actually a monoid, giving us that any element of κ is also a monoid in the category of endofunctors and thus $\Gamma \times \kappa = \emptyset$. In this case, though, the problem cannot possibly be less gross than this observation, giving us that $\mathbb{G}(P) \geq \text{"glasgow"}$.

3 Voodoo

In this paper, we will be concerned mostly with the lower bound of grossness as established via another measure we have provisionally named *voodoo*, denoted with $\mathcal{V}(P)$ defined as follows:

Take a problem $P \in \mathbb{E}$. If P can be reduced to some P' via an application of some subroutine Q , $\mathcal{V}(P) \geq n \cdot \mathbb{G}(P')$. It naturally falls out, then, that $\mathcal{V}(P) = \sum_{P \text{ reduces to } P'} \mathbb{G}(P')$. From here,

[†]This is the difference between the highest defined unicode codepoint and 95

we can apply the Rauen Incompleteness Principle to see that \mathcal{V} is a complete but undefined function.

From this, however, we can establish the following lemma:

Lemma 1 (Wong’s Lemma). *For all $P \in \mathbb{E}$, let L be the upper bound of $\mathbb{G}(P)$ as determined by the character counting method. Additionally, define $\mathfrak{D} = \mathbb{G}(P')$, where P' is the derivative of P . Then,*

$$\mathbb{G}(P) \geq \mathcal{V}(P)^{\mathfrak{D}} \pmod{L} \quad (1)$$

Proof. Fix a problem P . $\mathbb{G}(P')$ is trivially bounded from above by 1114016[†] and below by \mathbb{N} . We can then express $\mathcal{V}(P)$ type-theoretically as a combinatorial 4-form, which means that, by the Chinese remainder theorem, $\mathcal{V}(P)^{\mathfrak{D}}$ cannot be greater than a type-theoretic solution to P modulo L . But wait, we already established that $\mathbb{G}(P)$ is bounded from below by the set of type-theoretic solutions to P , so the claim immediately follows. \square

Note that this does not contradict our earlier statement about character-counting; the modulo ensures that we remain under our upper bound.

We now present the following result:

Theorem 1 (Reed’s Theorem). *For all $P, Q \in \mathbb{E}$,*

$$\mathbb{G}(P) \circ \mathbb{G}(Q) \geq \frac{\mathcal{V}(\overrightarrow{P \cdot \mathbb{G}(Q)} \times \overrightarrow{\mathbb{G}(P) \cdot Q})}{\mathcal{V}(P \circ Q)^{\mathfrak{D}_Q}} \quad (2)$$

where \circ is the composition operator.

Proof. We begin by applying Wong’s Lemma to the right hand side of the equation.

$$\frac{\frac{\mathcal{V}(\overrightarrow{P \cdot \mathbb{G}(Q)} \times \overrightarrow{\mathbb{G}(P) \cdot Q})}{\mathcal{V}(P \circ Q)^{\mathfrak{D}_Q}} \geq}{\frac{\mathfrak{D}_{\overrightarrow{P \cdot \mathbb{G}(Q)} \times \overrightarrow{\mathbb{G}(P) \cdot Q}} \mathbb{G}(\overrightarrow{P \cdot \mathbb{G}(Q)} \times \overrightarrow{\mathbb{G}(P) \cdot Q})}{\mathbb{G}P \circ Q \circ P}} \pmod{L}$$

Note that, at this point, all operators are now all near-computable with sufficient paper (and we can thus assume that our human prover has done so).

However, on attempting to apply the same thing to the left hand side, we get that

$$\mathbb{G}(P) \circ \mathbb{G}(Q) \leq \mathcal{V}(P)^{\mathfrak{D}_P} \circ \mathcal{V}(Q)^{\mathfrak{D}_Q} > (\mathcal{V}(P) \circ \mathcal{V}(Q))^{\mathfrak{D}_{P \circ Q}}$$

From here, however, the authors of this paper object (see section 2) to showing the remaining computations. Without lots of Generality,

we see that the two quantities thus have well-ordered grossness. Specifically, something that is near-computable cannot be greater than anything that is not. Because the LHS is ultimately too gross to reduce, then, it must be greater than the idealized RHS. \square

Because of this, we can also establish the following corollary

Corollary 1 (Falk Corollary). $(\mathbb{E}, \mathbb{G}, \circ)$ is a well-defined field.

4 The Principle of Repugnant Exclusion

Consider the 3-Body Problem (\cdot) . Notice that, despite being somewhat simple to state, it is difficult to intuit a value of $\mathbb{G}(\cdot)$.

Suppose P is as follows: *Given $\perp \Rightarrow lv$, find lv* . Simply enough, $\mathbb{G}(P) = \perp$. Note that P has triviality and \cdot has pure nontriviality, so we can apply the law of excluded middle to see that $\mathbb{G}(\cdot) = \top$. We will refer to this relation by saying that \cdot is *excluded* from P . Without loss of generality, we can generalize this proof to say that, given a problem P that can be excluded from a problem Q such that $\mathbb{G}(Q) = \perp$ must have $\mathbb{G}(P) = \top$ (and vice versa). This is the Principle of Repugnant Exclusion.

5 Objectability

We end by opening a discussion on *objectability*, the discussion of whether a problem is *objec-*

tionable. Consider the 3-Body Problem used to establish the Principle of Repugnant Exclusion. This problem involves several abstractions, reductions, equations and several calculus and algebraic impossibilities[‡]. It would be reasonable to say that no reasonable person could sit down and solve this problem, so we would call this problem *objectionable*. We consider this to be a simple enough definition with profound consequences. Formally, we believe that any problem P for which $\mathbb{G}(P)$ is greater than $|\langle P \rangle|^{L(P)}$ is objectionable, but we do not have a proof at the current time (we offer this as an open problem).

By way of example, consider the problem *Given a problem P , is P objectionable?* (this is the problem O). By applying earlier results from the Rauen Incompleteness Principle and Wong's Lemma, we see that O is undecidable. Any resulting decider M , then, cannot be encoded via any finite alphabet Σ , which means that it must contain an infinite number of non-common English symbols (thus immediately failing the character-counting test). Furthermore, attempting to analyze the types of $L(O)$ (the set of $\langle P \rangle$ for all P that are objectionable) is objectionable without a well-defined proof of the objectionability of O (see citation 2). If such a proof existed, then, it must itself be objectionable (and so we cannot assume that the proof was ever written or read by any human author). We, as humans, must thus conclude that O itself is objectionable.

References

1. A. M. Turing. On computable numbers, with an application to the entscheidungsproblem. Proceedings of the London Mathematical Society, s2-42(1):230–265, 1937
2. You thought there would be a citation, but it was me, Dio Brando!
3. “Anime was a mistake” - Hayao Miyazaki

[‡]It is also already unsolvable.

Towards A Well-Defined and Secure Flirtation Protocol

Rowan Copley,^{1*}

¹Department of Love Quantification, Witty Pear LLC

*To whom correspondence should be addressed: rowan.copley@gmail.com.

Introduction

Human interactions are inconsistent, ambiguous, and fraught with danger. Protocols for interaction are so nebulous that there is an entire cottage industry devoted to after-the-fact interpretations of interpersonal interactions[1]. In fact, many people will live their entire lives with their own internal understanding of proper communication protocols, which are mutually incompatible with other such implicitly defined systems.

This is especially true in the communication between sexes, frequently referred to in the vernacular as "flirtation." We believe that this is an unsatisfactory state of affairs that can be improved with the application of networking and information theory. Therefore we are proposing a protocol for inter-gender communication.

1 Architecture

Consider social encounters to be a nested stack of "consent". We start by defining a stack of consent levels that are traversed during the course of the flirtation session. To keep the definition

Table 1: The consent hierarchy.

Consent Level	Description
-2	Don't even look at me
-1	Don't talk to me
0	You may speak to me briefly if there's a good reason
1	We can talk
2	You can talk to me all you want
3	You can touch my hand
4	Long eye contact might not be creepy
...	How deep this stack goes is outside the scope of this paper.

intuitive and easy to understand, we define consent levels below in a colloquial and intentionally non-rigorous way.

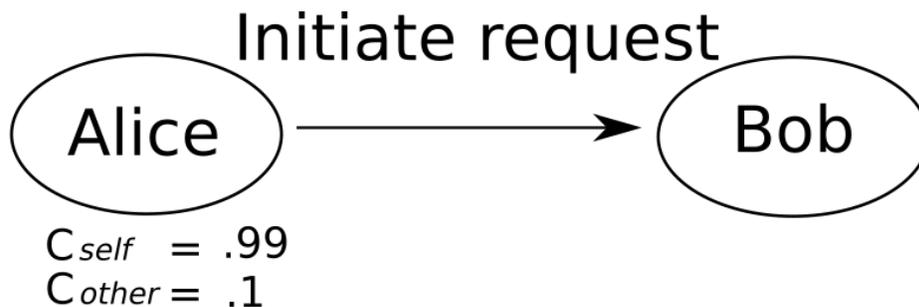
Behind this is the idea of a mutual consent to recurse on that stack to a "deeper" level. You'll need a protocol of mutual agreement to recurse that is also recursive: knowing that the other person knows that you know, and so on. But such a protocol can be implemented as a Two Generals' Problem, which, while technically unsolvable, is still provably simpler than everyday human interactions.

One design decision is whether consent levels should be symmetric, where each participant is cool with what is going on, or asymmetric, where maybe something creepy is going on. While our protocol is able to support both by letting the two parties negotiate this at runtime, we focus on the symmetric use-case here as (we hope) it is more common.

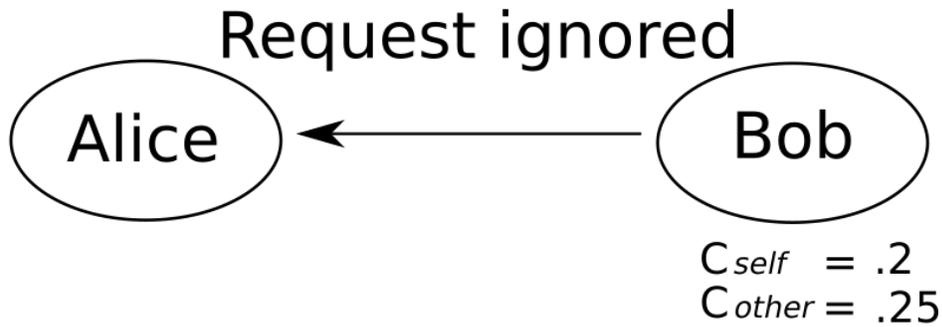
The heart of the protocol is each party's location in the stack, and the messages about accepting or denying access to deeper levels, as well as revoking access to a level that was previously accepted. Revoking access to a consent level is left up to the particular implementer, will depend on localized and internationalized norms. Some examples of revoking access are getting up and walking out, saying "no thanks", and kicking them in the balls.

2 Demonstration

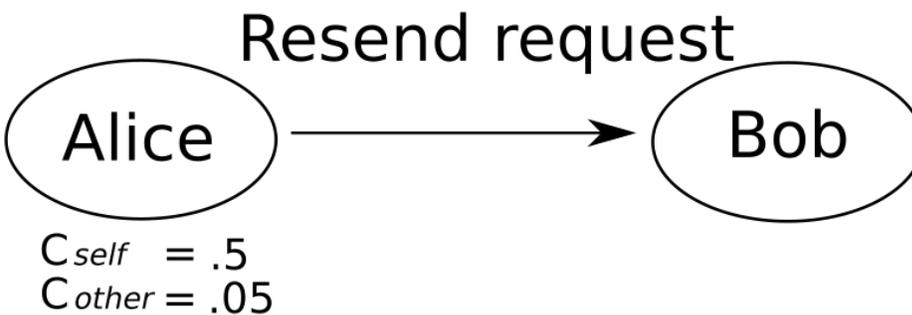
Suppose that Alice wishes to initiate a flirtation request with Bob. That is, she wishes to enter into a negotiated descent down the consent stack. Currently, they are only on speaking terms, but Alice wants more. However, Alice does not wish to be in a more consent level of the stack than Bob. That would expose Alice to embarrassment should Bob not be aware of Alice's intentions and similarly change positions on the stack. Alice has lead a life of disillusionment and disappointment which has left her with a calloused emotional exterior. And she's too old for pussy-footing around anymore. She's had it with immature young men whose only weapon is flattery. Alice wants a man who can be honest about his flaws, who opens up to you. She want Bob. Therefore, Alice and Bob must engage in a negotiation which proceeds as follows.



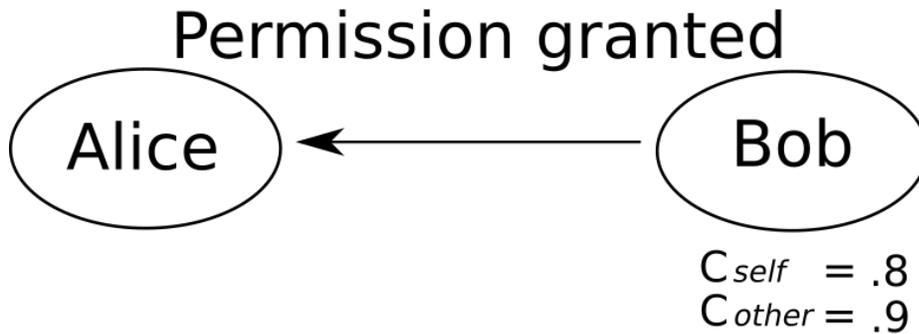
We define two variables, C_{self} and C_{other} , to indicate the degree of confidence that Alice in herself has that she wants to intimate a more flirtatious atmosphere with Bob (C_{self}) and the degree of confidence that Alice has in Bob that he wishes for the same.



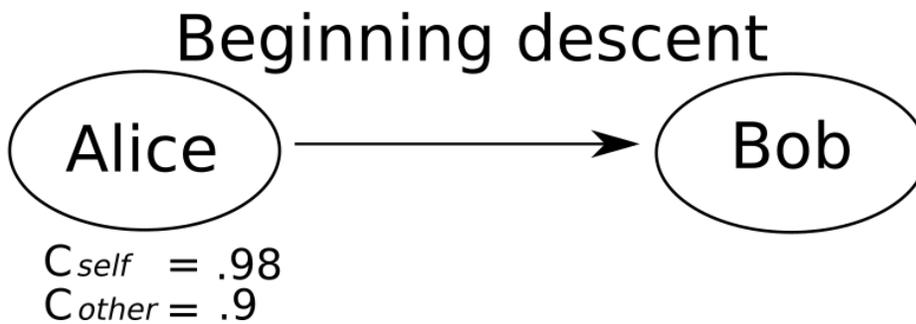
Here we see Bob's internal variables set at near their default levels. Because he has been oblivious to Alice's advances, he has no idea what her true intentions are and thus his internal model is quite inaccurate. The social engagement blunders onward, however, driven by Alice's determination.



We see that Alice has updated likelihood estimators to reflect a decrease in confidence that Bob wishes to engage in a flirtation exercise with her. However, Alice decides to try again by calculating that although the chances of payoff are estimated to be low, the penalty for failure is also low.



In this case, Bob makes the decision to initiate the simultaneous stack traversal immediately. This may not always be the case, as sometimes a delay is a safer option.



Alice takes the potentially risky decision to initiate a descent without receiving a response from Bob. Unfortunately, Bob interprets this as desperation and revokes permission to proceed at the last minute. Alice must bury her feelings deep inside her and only indirectly express them, such as in passive-aggressive post-it notes or clinical dissections of human behaviour in academic journals.

During this exchange, suppose Eve wishes to intercept Alice's flirtation message as if it were intended for her. Alice did not intend that Eve receive her obfuscated message to Bob. Alice desires only Bob, and considers Eve to be an obstreperous harlot. Alice's only wish is to hold Bob close, to feel his caresses.

Unfortunately, unlike in computing where each machine is assigned a convenient protocol address (e.g., 127.0.0.1), humans have no such unambiguous addressing. Furthermore, in a social setting where one is broadcasting flirtatious messages, simply claiming to be the recipient of those messages may lead to a self-fulfilling prophecy. Therefore, at least until such a time as this promising academic direction can lead to more concrete results, we recommend three potential solutions to this problem:

1. Distance method: always be the closest to the person for whom your flirtatious message is intended,
2. Focus method: make eye contact with that person continuously through the course of the social event,
3. Name method: use their name in every sentence.

For redundancy, using more than one of these methods will increase the likelihood of success.

3 Issues With Adoption

An important aspect to consider for any protocol is how difficult it will be to get a substantial percentage of the population to adopt it. While it is true that the protocol is not useful unless both parties have adopted it, in this case, we believe the advantages to the protocol are obvious and the rollout will be smooth. Furthermore, due to network effects, adoption will follow an exponential growth curve.

4 Future Work

We plan to expand the protocol to include flirtation sessions involving more than two participants, and are recruiting graduate and undergraduate research assistants.

References and Notes

1. Psychology. Wikipedia: the free encyclopedia

A Solution to the Two-Body Problem

Rose Bohrer

Abstract

The two-body problem asks whether two massive academics undergoing mutual attraction can locate stable employment and lodging. We identify an important special case in which a solution exists, the first of its kind. We give a lower bound on net worth as a function of time. Given the significance of the result, we give a near-formal proof in Differential Dynamic Logic $d\mathcal{L}$ to increase confidence in the result, without the hard work of actually doing the proof correctly.

1 Introduction

The three-body problem has been known since Newton: Given three massive bodies mutually exerting gravity, compute their trajectories. It has also been known since the time of Poincaré that no analytic closed-form solution exists in the general case. Poincaré was not discouraged by this. Rather, (mostly because he and his wife were both on the academic job market at the same time), he did what any good mathematician would do and simplified the problem until there was some hope of solvability. Poincaré’s simplification is what we now know as the *two-body problem*: “Given two massive academics in mutual attraction, determine stable employment and lodging”. This simplification is of massive importance to academics even today, and as the modern academic knows, no general solution has yet been found.

2 Related Work and Previous Solutions

In extremely special cases, satisfactory solutions to the two-body problem have been found. The most notable special case is that of Blum and Blum who have found simultaneous full-professorships at a Research I university in an affordable city in related research areas. This is, clearly, the most difficult instance of the 2-body problem. However, the B^2 solution requires at least one Turing award, and thus does not scale to us mere mortals. It would have worked for Poincaré, I’m sure, but not for me. Who am I even kidding, the 1-body problem has long been solved.

Other approaches are more ambitious, seeking to provide an answer to us mere mortals by sidestepping the requirement for an analytic closed-form solution, satisfying themselves with implicit or non-analytic solutions instead. The most notable approach is that of the Klein Four group, who take on the even more general three-body problem using nonanalytic functions such as a cappella: https://www.youtube.com/watch?v=Aiq_oaIvyak

3 Our Approach

The key issue here, of course, is to identify a salary for two PhD's at the same time, as only so many employers are willing to employ the unemployable. In offline dating, it has been widely recognized that the problem becomes significantly easier as the distance in thesis topics increases. Thus, the Two-Body problem can be reduced to the Some-Body Problem as posed by Mercury: **Can anybody find me somebody to love? Who has a PhD in a technical field, but preferably not CS and definitely not formal methods, oh god please not formal methods?**

Our approach is centrally based on Mercury's reduction. Further, we make the simplifying assumption that the relationship between the bodies is *stable*, a common assumption which significantly simplifies the dynamics. Because we provide the first and only solution for such a seminal problem in bodyology, we provide a formal model and proof of our solution to the two-body problem in Differential Dynamic Logic, an established logic for verifying hybrid systems, many of which are Cyber-Physical Systems. This marks its first known use in verifying Cyber-Social-Physical;-)-Academic Systems.

3.1 Model of Academics in Motion

In the two-body problem, we are given two rigid ;-)- romantic bodies in motion and seek to show they can sustain some minimum level of income over an extended period of time. As in the typical presentation of the problem, we assume the sole source of income is a university.

In this model, we consider two-dimensional rigid bodies, whose positions are described by x_1, y_1, x_2, y_2 . Following standard simplifications, we assume the position of the university is somewhere between the academics in a way we will make precise soon. According to government statistics, the salary for each academic is inversely proportional to the distance from the university. We add another variable t to track the evolution of time. Putting this all together, we arrive at a 10-dimensional ODE describing the evolution of the system:

$$\alpha \equiv \left\{ \begin{aligned} x'_1 &= v_{x,1}, v'_{x,1} = -\frac{m_2 \cdot v_{y,1}}{d^2}, x'_2 = v_{x,2}, v'_{x,2} = -\frac{m_1 \cdot v_{y,2}}{d^2}, \\ y'_1 &= v_{y,1}, v'_{y,1} = \frac{m_2 \cdot v_{x,1}}{d^2}, y'_2 = v_{y,2}, v'_{y,2} = -\frac{m_1 \cdot v_{x,2}}{d^2}, \\ \$' &= \frac{1}{d_{1,U}} + \frac{1}{d_{2,U}}, t' = 1 \end{aligned} \right\}$$

Where we define $d \equiv \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ and $d_{1,U} \equiv \sqrt{(x_1 - x_U)^2 + (y_1 - y_U)^2}$ and $d_{2,U} \equiv \sqrt{(x_2 - x_U)^2 + (y_2 - y_U)^2}$ where the university is at a fixed location x_U, y_U .

Like all good proofs we have some preconditions. Specifically, the mass and distance are in harmony such that the bodies will maintain a stable

orbit, and the university lie somewhere inside that orbit:

$$\begin{aligned} Pre &\equiv v_{x,1}^2 + v_{y,1}^2 = v_{x,2}^2 + v_{y,2}^2 \\ &= \frac{m_1}{d} = \frac{m_2}{d} \wedge t = 0 \\ &\wedge (x_U - (x_1 + x_2)/2)^2 + (y_U - (y_1 + y_2)/2)^2 \leq d^2 \end{aligned}$$

And our postcondition establishes a bound on income:

$$Post \equiv \$ \geq \$_0 + \frac{2 \cdot t}{d}$$

Putting these together, we arrive at our theorem-statement: **Stable Relationship Solution to the 2-Body Problem:**

$$Pre \rightarrow [\alpha]Post$$

Proof. We decompose the proof into its key lemmas, leaving the details of the proof as an exercise for the reader.

Lemma: Bounded Attraction. A stable relationship requires that the bodies be attracted to each other, but not so attracted to each other that they might lose their wits, or in our case, their orbits. The Bounded Attraction lemma asserts an upper bound on the attraction, or force:

$$[\alpha]v_{x,1}'^2 + v_{y,1}'^2 \leq \frac{m_2 \cdot v_1^2}{d^2}$$

And similar for the second body.

Lemma: Constant Separation. Next, we must prove the essential simplifying step that distance between the bodies is constant, without which the rest of the proof is intractible. Because distance makes the heart grow fonder, we call this lemma constant separation:

$$[\alpha]d' = 0$$

Lemma: Stable Relationship. Stability of relationships is essential to maintain the healthy life attitudes that allow one to succeed in academia. Thus we show the bodies are in a stable orbit at the radius $d/2$ which is centered around the origin, without loss of generality:

$$[\alpha]x_1^2 + y_1^2 = (d/2)^2$$

And so on for the other body.

Lemma: Home Sweet Home. In order to establish a bound on salary, we need a bound on the distance to the university. Because the university stays at its starting position inside the circle, we can bound this distance by the diameter:

$$[\alpha](x_1 - x_U)^2 + (y_1 - y_U)^2 \leq d^2$$

Lemma: Continuous Cash Flow These combine to give us a bound on the continuous rate at which cash changes, which leads directly to the theorem:

$$[\alpha]\$' \geq \frac{2 \cdot t}{d}$$

□

4 Conclusion

We have proven the first known lower bound on salary in the two-body problem under realistic situations. Practical applications of this solution abound for otherwise-broke academics on the job market. Our simple closed-form solution simplifies financial planning for the affected academic masses. We “formalized” the result for increased confidence, in case any... body should dispute its correctness.

Call For Partners: Romance with Rigor

Rose Bohrer

1 Introduction

Traditional dating sites are based on simple premise: participants are capable of identifying, with a reasonable degree of accuracy, which of their proposed matches are appropriate for them. Millions of years of empirical dating evidence show this premise to be false. As in all human pursuits, the dating field is rife with biases and irrational judgements, such as “You’re too ugly”, “you don’t make enough money”, and “at least he can remember my birthday, Rose.” While no system can completely eliminate human biases, the scientific community has, by and large, done an impeccable job of minimizing the influence of bias through its system of *peer review*. Even in non-blind review, simply entrusting the review process to a disinterested third party with no conflicts of interest greatly increases the quality of the outcome. If only the same could be said for love.

Unlike all other dating sites, `callfor.partners` addresses the *underlying* problem in online dating: lack of rigor. We do so by applying that most successful of human inventions, peer review. As with academic peer review, not all decisions are made by peers. Just like you get to choose what paper to submit, **YOU** are in charge of who you want to date, by writing your own personal **CALL FOR PARTNERS** and **YOU** get to make the best impression by submitting your own **PARTNERSHIP ABSTRACTS**. Only the messiest, most error-prone part is spread between peers: Deciding which partnerships to pursue and which to reject. The part that nobody wanted to do anyway.

2 Design

The central feature of the Call For Partners Partnering Workflow is the *Partnering Committee*. As with the peer review process, one selects a group of one’s closest friends to make decisions as to which advances should be accepted or rejected. Usually (but not always) the reviewing process is mutual: users are motivated to join their friends’ partnering committees for the reviewing services they themselves receive in return. The standard dating website trope of “profiles” appears in Call For Partners under the guise of Calls For Partners. The distinguishing feature of a CFP vs. the traditional profile is that unlike a traditional profile, a CFP is all about what **YOU** want, specifying in utmost precision the desired features in a partner. CFPs come in multiple styles. For example, a *Journal CFP* often has a rolling deadline or no deadline at all, where potential

partners are encouraged to submit at whatever time they find convenient. In contrast, a *Conference CFP* generally has a strict deadline which is extended only after a disappointingly small number of submissions. The Conference CFP is especially useful for implementing *rebound relationships*, a implemented on many sites, but never before with such rigor.

The CFP model acknowledges that you are TALENTED, and your history will speak for itself. In addition to a CFP, every participant has a CV listing relevant accomplishments. Because we are living in the future, Call For Partners applies advanced AI technology known as “Facebook Stalking” to automatically generate large portions of a CV. This open-access model (with a level of openness exceeding many leading-edge academic organizations including SIGPLAN) reduces harrassment and other abuses of the system, because participants are held accountable for their behavior in public. Remember, boys: **Before you do something you’ll regret, DBLP don’t forget.**

Given a CFP and CV, participants have all the information necessary to write a Partnership Abstract. A partnership abstract gives a brief, concise description of the contents of the proposed relationship. The Partnership Committee compares the abstracts against each author’s CV and own CFP, and uses this to write reviews, ranking each abstract and/or each individual’s self-worth on a scale of A-F. In most cases, the reviews are clear enough that the PC can reach an anonymous conclusion as to whether the abstract should be accepted or not. In the case of a dispute, a PC Chair can be appointed with tie-breaking authority.

Upon acceptance, the relevant parties gain the ability to message each other. Often, the author of the CFP will request revisions from the author of the Partnership Abstract before starting the Partnership. While the proposer does not have the ability to reject the proposee nor request changes of them, it is traditional to disclose further results publicly, which over time has a significant affect on the proposee’s Impact Factor.

As with most dating sites, CFP provides functionality to help you search through potential matches into to identify someone to whom you wish to submit a Partnership Abstract. In addition to the barebones necessities like salary, race, and number of previous partners, CFP allows you to perform custom searches that solve the problems specific to academic communities. In fact, we provide the first known solution to the Two-Body Problem, a generalization of the Three-Body Problem initially posed by Newton.

The Two-Body Problem was originally phrased by Newton is stated as follows: **Given two bodies A and B each with an attraction and a PhD, find a place of residency and a salary.**

The key issue here, of course, is to identify a salary for two PhD’s at the same time, as only so many employers are willing to employ the unemployable. In offline dating, it has been widely recognized that the problem becomes significantly easier as the distance in thesis topics increases. Thus, the Two-Body problem can be reduced to the Some-Body Problem as posed by Mercury: **Can anybody find me somebody to love? Who has a PhD in a technical field, but preferably not CS and definitely not formal methods, oh**

god please not formal methods?

At this point, the astute reader will notice that this problem is amenable to solution via an adequate search feature. The first-of-its-kind CFP Field Search allows one to specify the exact desired distance between their mate’s work and their own, using traditional metrics such as Er’os Numbers, Bacon Numbers, Erdos-Bacon Numbers, and also less traditional metrics. In our companion paper, we have verified our solution to the Two-Body Problem using Mercury’s Reduction.

3 Implementation

Call For Partners is currently available to the public as an open beta at `callfor.partners`. Call For Partners is implemented with Scala, Slick, Play, PostgreSQL, Heroku and assorted other buzzwords.

At present the implementation is limited, with the primary limitation being one known as “grad school”, specifically, “actual work”. However, and especially due to the subject matter, the implementation process has been utterly free of the other canonical time limitation known as “girlfriends,” which the author credits with the implementation progress made so far.

Producing a robust, production-ready implementation of an application like CFP requires many things. The most important requirement from a business perspective is the ability to rapidly acquire, and potentially disseminate to St. Petersburg, a wide variety of sensitive personal information on customers. In this critical area, we already accell, through a robust account application process. Our application process is based on well-established social engineering techniques that lull a user into a false sense of security before we go in for the kill. For example we ask them common questions such as their Social Security Number and mother’s maiden name before getting them to divulge information that many users find sensitive, such as their thesis topic.

Common wisdom in the software industry is that a successful web application also needs “features”. One contribution of our work is to refute the above, phony claim. Websites like Match.com and eHarmony.com have implemented almost all the features shown in their business proposals and an even more astonishing fraction of features featured in their advertising materials. CFP, being authored by the expeditious academics that it is, has more or less elided this step, yet derived an equal number of publications.

4 Evaluation

An early version of `callfor.partners` was tested on select attendees of SIBOVIK 2016 in a private beta for the past year. In retrospect, this choice of test users presented a huge logistical problem: In large part due to their collective scientific prowess, the average attendee of SIGBOVIK 2016 had at the start of our trial 2.3 girlfriends, 6 boyfriends, 1.9 wives, 4 husbands, 24 desperate exes trying to

get back together with them and the occasional 1.5 Texas. In order to keep the evaluation simple, all such relationships were terminated at the start of the study, the last of them much to the chagrin of James K. Polk. This in fact interfered significantly with the validity of our study by sending a vast tidal wave of fresh singles throughout the surrounding community. However, keeping with the long-held standards of the scientific community, we will publish the study anyway.

Our experimental evaluation sought to answer the following questions:

- How does CFP affect the overall number of people dating in a population? We call this the *mojo quotient*.
- How does CFP affect the romantic success of individuals within the population? We call this the *love-potion factor*.
- How effective is CFP at enabling the treatment group to restore itself to its equilibrium quantity of relationships when the equilibrium is disturbed? We call this the *rebound factor*.

A treatment group of 125 attendees was provided with an experimental pre-release version of CFP. A control group of 75 attendees was provided with industry-standard dating software. Out of the control group, only 15 participants made it to the end of the trial. Out of the remaining 60, 40 quit the trial early due to the despicable dating options available to them, 18 died of loneliness and two of them decided to date each other as an excuse to stop talking to hot singles in their area. By the end of the trial, the treatment group had reached a size of 500, for the participants had produced on average 3 offspring, with the remaining 25 additions consisting of a mixture of immaculate conceptions, Russian hackers, and Peruvian drug kingpins bribing their way into the trial.

We initially proposed that the mojo quotient be computed as the fraction of an overall population engaged in a romantic relationship, but this had two failings:

- The structure of the population changed vastly over the course of the study due to widespread procreation.
- By the end of the study, each participant had an average of 550 relationships, contracting common wisdom as to the value of Dunbar’s Number.

The former failing was ameliorated by the fact that a large number of the resulting newborns also started dating each other. However, this result was so surprising and potentially-unethical that we felt it, too, ought to be somehow reflected in our metric.

Thus we settled on the following definition of mojo quotient:

$$Q_M = \frac{\sum_{p \in \text{Participants}} |\text{rels}(p)| \cdot c^{\text{diapers}(p)}}{|\{p | \text{age}(p) \geq 18\}|}$$

Where rels is the number relationships a participant is engaged in, $\text{age}(p)$ is their age, $\text{diapers}(p)$ is the average number of diapers they use in a day, and c is an experimentally determined constant set to 5 for the purposes of this study.

The love-potion factor is defined as the median ratio of number of relationships/year after/before the introduction of CFP. Naturally this metric was computed only for participants already in existence at the beginning of the study. We computed the love-potion factor to be exactly 3, no more, no less, whereas the control group had a love-potion factor of 0.4. This factor was arguably not too meaningful given the large number of deserters.

The unique structure of this study greatly simplified our calculation of the rebound factor, because all participants were artificially reduced to 0 relationships at the onset of the study. The rebound factor is computed by calculating the time it takes for both the treatment group and control group to reach their initial fraction of romantically active participants, then taking the ratio between the two groups. Unfortunately, the control group never reached its initial fraction, and thus we have computed a rebound factor of ∞ .

In conclusion, CFP is infinitely better than all competing options.

5 Testimonials

As if the cold, hard, numbers from the previous section were not enough, we have also obtained a series of testimonial comments from participants in our study. While testimonials lack the rigorous proof of superiority that we already provided above, they aid us in the analysis of subjective aspects of the work, such as *why* we are the superior dating service.

Initially, I was skeptical about entrusting all my personal information to a website that encrypted my communications with a Vic cypher, and whose idea of a salted hash was cannabis with a dash of sodium chloride, but the minute Anastassia SQL-injected her way into my financial records, she injected herself into my heart! Nothing spells true love quite like breaking past a firewall, compromising private records and selling my credit card to the Russian mafia. We've been together for months now and we really just click. — Ryan Kavanagh

Before I used CFP, I was in, like, one relationship, tops. Now I'm in so many relationships I think the US Navy is jealous of just how often I am shipped. — Stefan Muller

6 Related Work

Many websites have addressed the related work of dating, though the author strongly suggest that you do not date your relatives for it is bad for the gene pool. Websites such as eHarmony are based on the pseudo-science of compatibility. Anyone who has ever tried to open a Word 2016 in Word 95 knows that most

things we say are compatible, are not, and thus it is with humans. OkCupid appeals to paganistic rituals in the hopes of receiving optimal pairings from the divines, who, unfortunately, do not believe in computers. Match.com and Tinder both determine optimal dates by lighting large groups of singles on fire and seeing which ones burn to the ground. Those which do not burn are witches and thus not good dating material, but unfortunately by that point all the good ones are dead, making the method ineffective in practice.

7 Promotional Offers

Most scientific papers offer you so-called “knowledge” for “free”, where “free” means you already belong to an institution which has already paid an exhorbinant fee in order to access the publication, or perhaps you have located the author’s so-called “web-site”. At CFP, we believe in doing things differently. This article has already provided you with free knowledge, free not only as in beer but as in speech, and now we are going to provide you with even more freedom like the good patriotic Americans that we are.

Typically, like all profitable businesses, CFP does not come for free. It can be paid for in several different ways, including not only the typical monthly subscription models, but also more academic-friendly methods such as co-author status on your publications or NSF grant funding (after extensive lobbying of the NSF Computer Science Directorate, acquisition of life partners is now considered a billable research expense). Since you, dear reader, have gotten in early by reading our very first publication, you can try CFP for a limited time, free-of-cost. This year’s SIGBOVIK 2017 proceedings come with an included 6-month membership to CFP (this was totally not a bribe in order to get the present article published), which can be extended to at least 2 years via our affiliate program: every friend or lover you recommend to CFP (to the paid service, of course) will extend your membership by 1 month.

Now how’s THAT for some science?

8 Conclusion

In this paper, we have presented Call For Partners, an online dating service based on academic peer-review. Call For Partners uses the rigor of peer-review to reduce the number of dating errors due to human bias. An implementation is provided using the Inter-Net, by which users can try out the proposed dating methods for themselves and significantly increase the author’s material wealth. An empirical study shows that not only does CFP significantly reduce the number of errors, but it drastically increases the amount of dating at the same time. Not only are the results of our empirical study significant, but a promotional offer is given whose savings are significant as well.

Nano Electronic Transplant in Brain and Eyes to Analyze and Process Human Thoughts and Emotions.

Chinmaya Lele

University of Pittsburgh

chinmaylele1993@gmail.com

Striking a conversation is very important and saying the right words is as important as to identify what's going in the mind of the other person. Identifying the emotions of the other person helps us in narrowing down the things which we must say to turn the conversations into opportunities.

The transplant is supposed to be implanted in brain and eyes, where the eye implant would capture the facial expressions of the person. The brain implant is going to communicate with the eye implant which would send the information signals to the brain implant to analyze the information captured.

The brain implant will analyze the information which would send back signals to the eye implant to display the appropriate response to be given to the person while engaged in a conversation. The eye implant would also have a omni display interface which helps the implant wearing person to see the information displayed. The eye implant is of the size of an eye contact lens which can be worn out when not required and which is charged through the solar energy which the eye absorbs while looking around. The brain implant would also have space to store the conversations which are being recorded by the eye implant which could be retrieved to be seen again and reviewed on the eye implant.

The idea resembles to the 2008 movie 'Iron Man' in which Tony Stark could see all the information through his display inside the armor suit. The only difference is that in this case it is the implant which is doing all the processing of images and emotions.

Grant proposal: Monetary policy of sparkly things

Pete and Luna, co-purrnciple investigators

Stefan Muller, research assistant

1 Introduction to sparkly things

We propose an investigation into the policy surrounding a currency recently introduced into our economy, known as the sparkly thing.



Figure 1: The purrnciple investigators with the sparkly things.

Sparkly things are earned by individuals as compensation for labor (see Figure 2) and can be exchanged for products and services such as prime lounging spots (see Figure 3).



Figure 2: Working hard to earn sparkly things.



Figure 3: The spoils purchased with sparkly things.

Sparkly things can also be saved in long-term accounts.

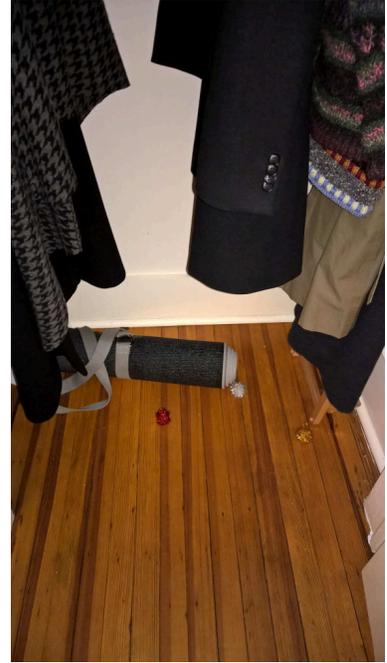


Figure 4: Savings account.

2 Central bank of sparkly things

We believe that the supply of sparkly things is controlled by the Sparkly Thing Reserve System, headquartered in the silverware drawer (Figure 5).



Figure 5: The Sparkly Thing Reserve Bank.

3 Interest Rates

The interest rate was recently lowered for the third time in two months, after we discovered that the sparkly things are not food.

4 Financial Crimes

As part of our investigation, we hope to discover ways of better identifying crimes such as money laundering.





Monkey track

It's Only a Game Theory



- 10 **Is it Percival time yet?: A preliminary analysis of Avalon gameplay and strategy**

Yuzuko Nakamura

- 11 **Dr. Boozehead, or How I learned to stop worrying and get drunk: Design principles and analysis of drinking games in the silicon age**

Kelvin M. Liu-Huang and Emily J. Simon

- 12 **A boring follow-up paper to “Which ITG stepcharts are turniest?” titled, “Which ITG stepcharts are crossoveriest and/or footswitchiest?”**

Ben Blum

Is it Percival time yet?: A preliminary analysis of Avalon gameplay and strategy

Yuzuko Nakamura
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213
yuzi@cs.cmu.edu

ABSTRACT

The Resistance: Avalon is a hidden-roles-style board game. In this paper, we use data collected over dozens of Avalon games to make recommendations on role sets and game sizes that maximize the game-playing experience. We also evaluate the effect of various strategies on good and evil's chances of winning.

KEYWORDS

Board games, hidden-role games, game design

ACM Reference format:

Yuzuko Nakamura. 2017. Is it Percival time yet?: A preliminary analysis of Avalon gameplay and strategy. In *Proceedings of SIGBOVIK, Pittsburgh, PA USA, April 2017 (SIGBOVIK'17)*, 6 pages.
DOI: 10.475/123_4

1 INTRODUCTION

The Resistance: Avalon [1], like Mafia, is a multiplayer game centered around hidden roles. Hidden role games involve players being randomly assigned roles that are not revealed to other players. These games often feature two or more sides with their own win conditions; in particular, there is frequently an evil or sabotaging side that attempts to bluff and win people's trust in order to win the game, and a good (but generally information-less) side that must correctly guess who to trust in order to win the game.

Avalon is a two-team game themed around King Arthur: the loyal knights of King Arthur (good team) attempt to succeed three quests (missions), and the minions of Mordred (evil team) attempt to be placed on the missions so as to sabotage them and lead three of them to fail. As such, the aim of evil is to be trusted by good players and the aim of good is to determine who can be safely trusted to be sent on a mission. In addition, all members of the evil team know each other (with possible exceptions).

Avalon in its simplest form¹ features two special roles, Merlin and the Assassin. The addition of two other special roles, Percival and Morgana, creates more opportunities for bluffing, trust, and strategy so we are interested in games with these four roles:

¹Avalon may also be played without any special roles (in which case it resembles its predecessor, The Resistance). However, the main feature of Avalon are these special roles.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

SIGBOVIK'17, Pittsburgh, PA USA

© 2017 Copyright held by the owner/author(s). 123-4567-24-567/08/06...\$15.00
DOI: 10.475/123_4

- Merlin (good): The player with this role knows all evil roles. They must guide other good players to this knowledge. However, Merlin must be subtle in the way they do this due to the role of the Assassin.
- Assassin (evil): If good successfully passes three missions, the player with this role gets to choose one good person to assassinate (after discussing with the rest of the evil team). If the assassinated player was Merlin, evil snatches victory from the jaws of defeat and wins the game. As such, this person serves as a check on Merlin's ability to help the good team.
- Percival (good): The player with this role knows who Merlin is. As such, they can also gain indirect information about who to trust by quietly observing Merlin's actions and can appear to evil as a decoy Merlin. However, due to the role of Morgana, Percival must first determine which Merlin to trust.
- Morgana (evil): The player with this role appears as a second Merlin to Percival, forcing Percival to spend some time determining who to trust, and possibly leading Percival to sabotage the good team by trusting Morgana instead of Merlin.

Good players who are neither Merlin nor Percival (generic good) know nothing about any player other than themselves. Merlin and evil players know the full evil team. Percival knows the two people who are Merlin and Morgana, but not which is which. This selective information is revealed to the appropriate players at the start of the game - termed the "nighttime phase" - by asking players to close their eyes, and then open their eyes to get information or raise their thumb to identify themselves, as appropriate.

In this paper, we seek to evaluate how rules and game size affect the funness of the gaming experience, and how different strategies are more or less successful for good and evil.

2 GAME ANALYSIS

2.1 Evaluating funness

2.1.1 Ideal win ratio. One feature of Avalon is that, not only are the good and evil teams asymmetrical (having different abilities, information, and objectives), but they are also of asymmetric sizes (the evil team in Mafia-style games needing to be a minority to avoid the game being trivially easy).

Supposition A: Funness is maximized by maximizing uncertainty. In other words, we want the probability of winning to be 1/2 regardless which team one is on. This means overall win chance of

Table 1: Ideal win ratios for each possible size of Avalon game under Supposition B (making evil wins as likely as good wins).

Game Size	# Evil	Ideal Good Win Chance
5	2	.40
6	2	.33
7	3	.43
8	3	.38
9	3	.33
10	4	.40

good and evil must be balanced to be 1/2 for an ideal game-playing experience.

Supposition B: Funness is maximized by making each person's wins equally likely to be earned while on the good team as while on the evil team. Equivalently, across all games, the good and evil teams both produce roughly the same number of winners. Under this supposition:

$$Pr(i \text{ is good} \mid i \text{ wins}) = Pr(i \text{ is evil} \mid i \text{ wins}) = \frac{1}{2} \quad (1)$$

Or...

$$\frac{Pr(i \text{ is good} \cap i \text{ wins})}{Pr(i \text{ wins})} = \frac{Pr(i \text{ is evil} \cap i \text{ wins})}{Pr(i \text{ wins})} \quad (2)$$

Assuming player i doesn't affect the win probability of their team, this is the same as:

$$Pr(i \text{ is good}) \cdot Pr(\text{good wins}) = Pr(i \text{ is evil}) \cdot Pr(\text{evil wins}) \quad (3)$$

If p_{good} is the win probability for good, and G is the chance of being good (i.e. the number of good roles over the total game size), then this equation is:

$$G \cdot p_{good} = (1 - G) \cdot (1 - p_{good}) \quad (4)$$

$$G \cdot p_{good} = 1 - G - p_{good} + G \cdot p_{good} \quad (5)$$

$$p_{good} = 1 - G \quad (6)$$

In other words, under this supposition of equalizing the portion of good and evil wins, the ideal win ratio for win is $1 - G$, or the chance of being evil.

The number of evil players changes depending on the total number of players. Table 1 summarizes the ideal win chances under this second supposition.

2.1.2 Game duration. Another important component of fun is the length of a board game session. We hypothesize that length of game goes up as the number of players in the game increases due to more discussion. We also are interested in comparing the typical game length to the 30 minutes claimed on the box.

2.2 Strategy

2.2.1 Percival claims. The rules in the instruction manual are not clear on whether players are allowed to claim to be Percival. However, the role of Percival is similar to the role of generic good – and unlike Merlin or the evil roles – in that claiming the role can potentially help the claimant's team (whether good or evil). Therefore, we allow players to publicly claim to be Percival.

A true Percival claim (Percival claiming Percival) can increase trust among good members but can possibly make Merlin assassination easier for the evil team. We are interested in whether claiming to be Percival, and the timing of such claims, tends to help good or evil.

2.2.2 The first mission fail. Evil players have the choice whether to throw in a fail card or a success card for missions that they go on. A sole evil player on the first mission may decide to pass the mission to avoid detection / suspicion for being on a failing team. However, an early mission fail can make the evil team's task of failing three missions total easier.

We are interested in whether first mission fails overall help the good or evil team, and how the size of the first mission factors in to this.

2.2.3 Evil coordination failures. When two or more evil players are on a mission team, they each have to choose whether to throw in a fail or success card, not knowing what their teammates are planning to do. As a result, sometimes evil players may end up passing the mission, or may put in more than one fail card, revealing key information about the make-up of the team. As such, a team with more than one evil person is not ideal for the evil team, and they may be cautious about proposing or approving teams with this make-up.

How often do coordination failures happen, and how do they affect evil's chance of winning? We investigate these questions in this paper.

3 METHOD

A body of 38 graduate students played games of Avalon (20 of which played "semi-regularly" i.e. five or more times during the data collection period). In total, 66 games were recorded although 5 were discarded due to incomplete information, resulting in 61 games overall.

All games were played using the Merlin, Percival, Morgana, and Assassin special roles. In addition, 10 of these games added the special roles Mordred and/or Oberon.

The following data were collected for each game:

- Number of players and role of each
- Approved mission teams and mission outcomes (but not proposed mission teams)
- Outcome of each mission (pass/fail)
- Outcome of game (win for either good or evil), including the win condition (three mission fails (evil win), mission success but Merlin assassination (evil win), or mission success and failed Merlin assassination (only good win condition))
- Which player(s) claimed to be Percival and when (if applicable)

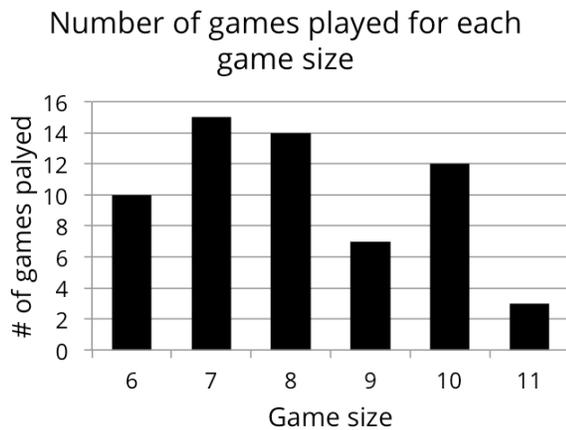


Figure 1: Number of games played for each size of game.

- Which player was assassinated (if applicable)
- Duration of game (measured from the end of night-time phase to either three mission fails or evil’s Merlin assassination choice)

Linear regression was used to determine whether game size affected duration.

Chi-squared tests were used to determine whether (1) the presence of Percival claims affected the evil’s team Merlin guess rate (Percival claim/no claim vs. Merlin guess/good win condition); (2) evil failing the first mission affected the chance of evil winning (first mission pass/fail vs. winner of game); (3) the presence of missions requiring evil coordination affected the chance of evil winning (zero/non-zero coordination missions in game vs. winner of game).

4 RESULTS

4.1 Number and size of games

Fig. 1 shows how many games of each size were played in the dataset. Although Avalon can in theory be played with game sizes of 5 to 10, players did not enjoy games of size 5 and so only played Avalon if at least 6 players were present. A game of size 11 can be played with the 10-player board and 4 evil characters (as in a 10-player game), and an extra set of vote tokens.

4.2 Win ratio

Overall, the win rate of the good team was .34. Fig. 2 shows how this win ratio changes with the size of the game. The good win ratio for 9-player games stands out as unusually high. This is also the game size with the fewest data points (see Fig. 1), so that may be part of the reason.

Under Supposition A of game funness, 6- and 9-player games are the only ones close to the ideal difficulty for good. 7-, 8-, and 10-player games fall short of both ideal win ratios. As such, it may be worthwhile to use gameplay mechanics that tilt the game in favor of good (Oberon as one of the evil roles, Lady of the Lake, etc.).

Under Supposition B of game funness, the 6- and 9-player games need to be altered to be more difficult. In particular, a 9-player

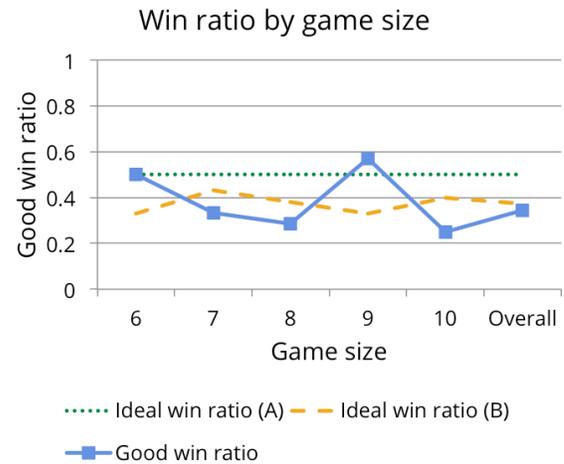


Figure 2: Good’s win record at each game size. The dotted line shows the ideal .5 win ratio under the supposition that good and evil should be equally likely to win. The dashed line shows the ideal win ratio under the supposition that people earn wins equally as good people as they do as evil people.

game might benefit from 4 evil roles (instead of 3), one of which is Oberon.

4.3 Game duration

Fig. 3 shows the distribution of game length. The mean game length is 57.3 minutes and the median game length is similar – 57 minutes. Most (80% of) games can be played within 80 minutes. This is markedly longer than the 30 minutes estimated in marketing materials.²

We can break down game length by the size of game, resulting in Fig. 4. Games of size 9 are again an outlier, being unusually quick, and being the only game size that approaches the 30-minute estimated play time.

There seems to a slight trend of longer games with more players in line with our hypothesis; however linear regression (removing the 9-person games) does not quite reach significance ($p=.0663$) and game size has low explanatory power for duration ($R^2=.0524$).

4.4 Percival claims

Fig. 5 compares the outcome of games where a Percival claim is made vs. ones where no Percival claims are made. Games with Percival claims are much more likely to end in mission failure, which makes sense because one reason why Percival might claim is because several failing missions have happened, and Percival (or Merlin) is trying to increase the chance of choosing an all-good team (i.e. scenarios with multiple failing missions are scenarios where Percival is likely to claim).

Among the remaining cases where three missions succeed, we are interested in comparing how often Merlin is assassinated in

²It is possible that this particular group of graduate students discusses an unusually large amount during Avalon.

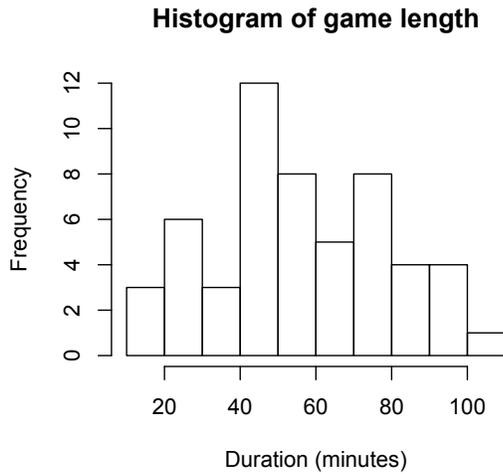


Figure 3: Histogram of game lengths.

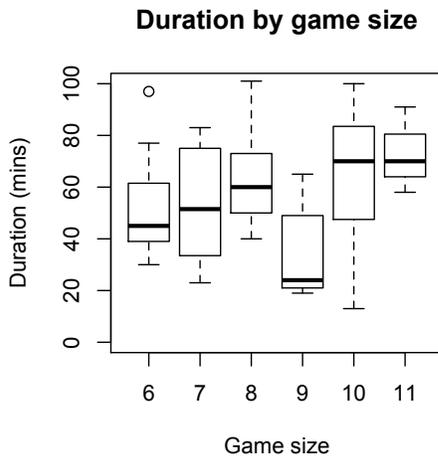


Figure 4: How length of game changes with game size.

Table 2: Merlin assassination successes under the conditions of a Percival claim vs. no such claim.

	Merlin assassinated	Good wins
Percival claim	16	11
No Percival claim	10	10

each case (Percival reveal vs. no reveal). Table 2 shows the number of games in each condition. The Merlin assassination chance when Percival reveals (.59) is higher than when there is no Percival reveal (.50). However, the chi-squared test shows no significant difference ($p=.738$).

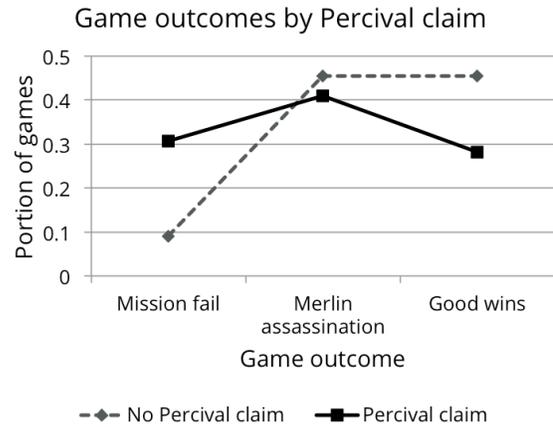


Figure 5: Portion of games that end with mission fails (evil win), Merlin assassinations (evil win), or neither (good win) under the conditions of Percival not revealing and Percival revealing.

Table 3: Game victor under the conditions of a first mission fail vs. a first mission pass (games sizes 8+).

	Evil	Good
First mission fail	13	3
First mission pass	5	4

There was not enough data to do an analysis of how the timing of Percival claims affect good's chance of victory. We leave this to future work.

4.5 The first mission fail

Although the intent was to analyze how first mission team size (two-person first missions (in games with 5-7 players) vs. three-person first missions (in games with 8+ players)) affects the outcome of the game, in practice, only one two-person first mission with an evil player (out of 16) was failed by the evil player. This is a difficult strategy to pull off for the evil player because for the rest of the game, at least one good person knows for sure one member of the evil team, and the evil person must consistently behave to give the impression of being someone in that situation.

Therefore, we instead look only at games with three-person first missions. Of the 24 games with evil players present on the first mission, 16 (67%) were failed by those players. Table 3 shows how evil's play during the first mission affected the victor of the game.

There is not enough data to perform a reliable chi-squared analysis, but it is possible that failing the first mission is overall a good strategy for the evil team.

4.6 Evil coordination failures

Of the 61 games, 32 (52%) featured no evil coordination missions, while the rest had at least one evil coordination team. Fig. 6 indicates how often games featured a certain number of evil coordination

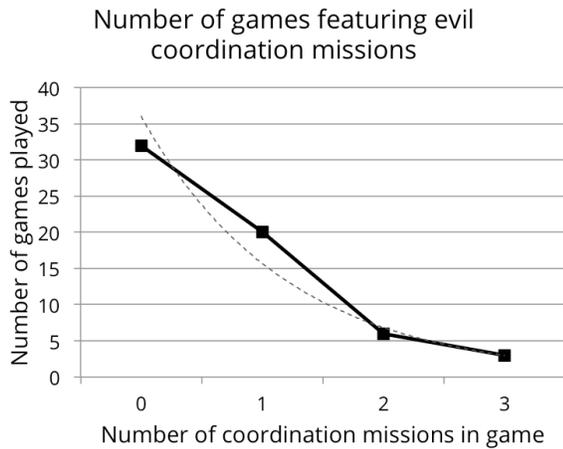


Figure 6: Number of games featuring zero, one, two, or three missions with more evil people than the required number of fails.

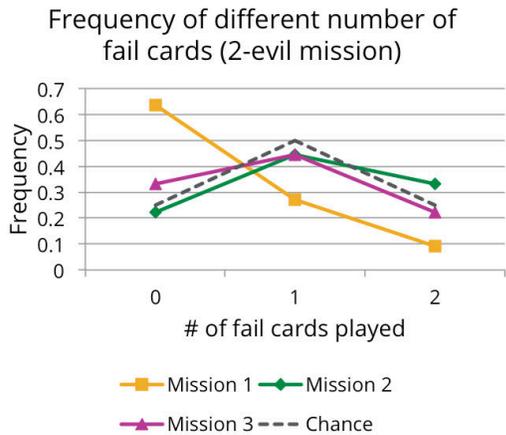


Figure 7: How often zero, one, or two fails come out for one-fail-required missions.

teams. Overall, this suggests the chance of any mission containing multiple evil people is roughly 15%. Note: It is impossible to have coordination issues on the fifth mission because any number of fails is acceptable for the evil team. Coordination issues on the fourth mission of games of 7+ players, where two fails are required, are rarer (as this only happens when three evil people are placed on the team) but are still important to the game.

38 of the 41 coordination missions (92%) involved two evil people on one-fail-required missions. Fig. 7 summarizes how frequently zero, one, or two fails come out in this situation. This figure shows that the number of fails that come out in Mission 2 and Mission 3 are roughly what you'd expect based on random chance (independent events with .5 probability of occurring). However, Mission 1 is much more skewed toward zero fails, corresponding to roughly a 20-25% chance of each person throwing out a fail. This makes sense as a two-fail result on the first mission can be costly to the evil team.

Table 4: Game victor under the conditions of zero or at least one evil coordination mission.

	Evil	Good
No evil coordination missions	15	10
1+ evil coordination mission	22	7

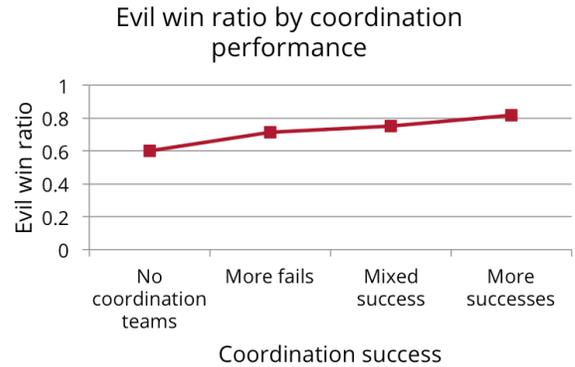


Figure 8: Evil win rate broken down by ability of evil to coordinate.

We also analyze how evil coordination missions affect the chance of good or evil winning the game. Removing from consideration games where evil never gets the chance to go on any mission, Table 4 summarizes the game outcomes when there are no evil coordination missions vs. when there's at least one. Evil's win ratio in the presence of coordination missions (.76) is higher than when there are no coordination missions (.60), although this effect does not reach significance ($p=.338$).

Figs. 8 and 9 break down the effect of evil coordination further. Fig. 8 takes into account evil's coordination performance – success means throwing out exactly the number of fails needed to fail the mission, and failure means throwing out more or fewer fails than needed. Fig. 8 separates the data into three conditions: games where evil mostly failed at coordinating (14 games), games where evil both succeeded and failed at coordinating once (4 games), and games where evil more often succeeded at coordinating (11 games). Fig. 9 shows how evil's win rate changed as the number of coordination missions in the game increased.

In all cases, there is not enough data to draw any definitive conclusions. However, contrary to expectations, it is possible that evil coordination situations might be slightly beneficial to the evil team.

5 DISCUSSION

We analyzed data from 61 games of Avalon. We found that games of size 9 were unusual in the amount they were played (less popular), how long they lasted (shorter), and game difficulty (good team more likely to win). Some games might require adjustment to their difficulty. In particular, 9-player games might require more evil characters, and 7-, 8-, and 10-player games might require a slight good handicap. Typical game time is more than one hour.



Figure 9: Evil win rate broken down by how many times during the game evil needed to coordinate.

In this particular dataset, Percival reveals resulted in slightly more Merlin assassinations; evil failing the first mission resulted in more evil wins; and the presence of evil coordination missions resulted in more evil wins. One possible explanation for this last finding is that it is hard for good people to reason about teams where more than one of the members was evil, and so they may be more likely to make decisions that assume only one evil person was on the team.³ However, more data might reveal these trends to be spurious/random noise.

A notable gap in the dataset is the general absence of two-person first missions failed by an evil player on the team. In the future, it would be interesting for evil to experiment with failing two-person missions to see if this strategy might be beneficial for evil overall. A more detailed analysis of Percival claim timing would be also be good to do with more data. Another promising avenue of future work would be to analyze the effect of rejecting missions on good and evil's chance of winning.

5.1 Conclusion

Hidden role games like Avalon provide a large space for both game design (e.g. number of players, set of specialized roles) and player strategy (e.g. failing the first mission, claiming Percival, team approval strategies, etc.). As such, collecting and analyzing data under different game conditions can be useful in improving the player experience and evaluating the strength of different strategies. Although limited by the amount of data, this work represents a preliminary step in the direction of analyzing the gameplay of Avalon.

ACKNOWLEDGMENTS

The author would like to thank Kristy Gardner for kick-starting this collective Avalon Problem in the department, Ryan Kavanagh for taking better care of the Avalon stats sheet than the author, and Liam K. Bright for listening to the author's shaky probability math.

³If evil does indeed benefit from coordination issues, this might interestingly increase the value of Zhu-Brown strategies (multiple evil proposing multiple evil people on their teams).

REFERENCES

- [1] Don Eskridge. 2012. *The Resistance: Avalon*. Indie Boards and Cards. Board game.



CONFIDENTIAL COMMITTEE MATERIALS

SIGBOVIK 2017 Paper Review

Paper 92: Is it Percival time yet?:
A preliminary analysis of Avalon
gameplay and strategy

Percival

Rating: Vote your conscience

Confidence: I'm so confused, guys...

It's no fun when Merlin gets assassinated or evil wins, so Supposition A sounds like something a spy would say. As a result, I question one of this paper's fundamental premises. Is the author a spy?! Regardless, this paper is a first-class analysis of our collective obsession with Avalon.

Percival II

Vote: Approve

Confidence: We've got to do this.

I'm totally not evil, and I'm the *real* Percival. I support this mission.

Ryan Kavanagh

Rating: Approve

Confidence: I think Percival is the real Percival?

No more Percivals, please! Think of the poor stats sheet!

Normally we would recuse ourselves from reviewing papers we're involved with,¹ but the program committee couldn't think of anybody anybody who hadn't contributed to this paper's data set. This paper provides a fun analysis of Avalon's innate funness. Section 4.5 claims there was insufficient data to perform a reliable chi-squared analysis. Consequently, I encourage the author to switch her research area to the analysis of Avalon and hire the program committee as research assistants so that we can play more Avalon, err, I mean, help her collect important data for a follow-up paper.

¹Because SIGBOVIK is a serious conference with serious reviews.

Dr. Boozehead, or How I Learned to Stop Worrying and Get Drunk: Design Principles and Analysis of Drinking Games in the Silicon Age

Kelvin M. Liu-Huang
Carnegie Mellon University
kmliu@cmu.edu

Emily J. Simon
Carnegie Mellon University
ejsimon@andrew.cmu.edu

Abstract

From beer pong to beer bong, drinking games have a storied past, seated at the intersection of sublimating puritanical repression [1] and the great ape's boundless curiosity. Animals utilize play to express themselves and practice behaviors. For humans, play is so important that rules of play are codified into games. Yet, scientific study of human games and game design has been greatly underrepresented, and even more so for drinking games. In the present study we sought to distill the essential principles of those traditions, which lie at the intersection of interactive gaming and indulging in poisonous fluids. Through careful field analysis and repetitive study, we propose that concrete prerequisites, mental requirements, and social abetment are all fundamental attributes of a successful drinking game. To evaluate our design principles, we designed three novel drinking games, beer baseball, soccer shots, and beer nim. We also evaluate the popular drinking game, beer pong, as a benchmark. Comparing our innovations to the benchmark, we demonstrate the effectiveness of applying our design principles, showing that beer baseball and beer pong knock it out of the park, while beer nim (our straw man) eats dirt.

1. Introduction

Animals evolved play to communicate and manipulate [2][3]; learn aggressive, predatory, and foraging behaviors [4]; and improve cognitive function [5]. To the great ape, play is so important that rules of play are codified into numerous philosophies called "games." Popular games are standardized internationally, generously funded, vicariously enjoyed by large fractions of the population, and game elders typically receive the

highest salaries at learning institutions [6]. Furthermore, games (as well as all other activities) are often integrated with ingestion of poisonous liquids to stimulate social interaction and enjoyment. In many ways, these "drinking games" may be regarded as paragon forms of play because they achieve so many different objectives of play.

Despite the importance of such games and the complexities of game design, very little formal study and scientific discourse have been devoted to game design. Ordinary tabletop games require delicate balance of tool complexity, rule complexity, computational complexity, game-to-game variance, audience appeal, and mechanical and narrative harmony.

The design of drinking games requires arguably even more sensitivity. Between the innately chaotic environment of parties, the need to facilitate communication, and judicious application of refreshments [1], drinking games embody the highest achievements of human design gathered from the likes of Chess, Go, or Ping Pong. Yet the design and study of drinking games is even lower in the pecking order than ordinary games. Even fewer serious examinations have been made of drinking game design [1][a][8]. Popular with men and women fond of classical languages, drinking games have historically been typecast as intellectually and socially inferior. At the risk of resorting to platitudes, we know that correlation does not imply causation [9], so this alone should not be an indictment of the noble pastime of drinking games.

2. Design Principles

A fecund party is a palpable maelstrom of active, bass/brainless, clumsy, dance, and entropy. Look

for these symptoms using the simple acronym, ABCDE. A drinking game should satisfy all the principles of game design, as well as judiciously accounting for these party properties.

2.1. Easy to organize (ABCDE)

Due to spontaneity and inattentiveness (A), a drinking game must require minimal planning, simple setup, and little infrastructure. Due to heavy bass (B), brainlessness (B), and entropy (E), mobilizing players and organizing a game must be simple. Props (if used) should be low cost and ubiquitous, or at least portable. There should be few and simple rules to explain due to (A) as well as interjection from the hard bassline (B). Due to brainlessness (B), clumsiness (C), and entropy (E), the drinking game should be low risk. Messes and injuries are sure to dampen a thriving party. Above all, the drinking game needs to be technically feasible given the specific parameters of the party. Space for a large game can depreciate due to dancing (D) and entropy (E). Too much bass (B) might also drown out the speaking portion of some would-be drinking games.

2.2. Social (NP, P=NP, KEG)

Parties must facilitate social interaction to avoid noncompliant prairie-dogs (NP), individuals who wallflower, stand alone, or look around perplexed. In general, we don't want players not playing (P=NP). A drinking game readily serves this need by providing a platform for players to communicate [1]. Meanwhile the game itself cannot require too much focus, so as to allow informal conversation. To facilitate social networking, a drinking game ideally allows players to join or leave as they please. We introduce a metric for this fungibility called the KEG (keep entering/exiting games) norm. Though some partygoers may wish to linger on one game, the option to devote only an aliquot of time is vital. Therefore we must always remember the KEG!

2.3. Appropriate difficulty (NP-complete)

The computational complexity of such critically acclaimed board games as Agricola and 7 Wonders tend to be unpalatable for a drinking environment. Other forms of play, such as football, hunting, and monster truck driving, carry a level of risk and finesse that should not be expected of inebriated patrons, due to brainlessness (B) and clumsiness (C). That is not to say that refreshments do not go well with a titillating round of Elder Dragon Highlander, but rather, the choice of drinking game depends heavily on the mood and flavor of the party.

Because intoxication impairs judgment (B), a drinking game has an ideal runtime complexity between $O(0)$ and that of ordinary games, inclusive. As with ordinary games, the level of difficulty needs to be carefully chosen, commensurate to the mood and audience. The game is boring if too easy and either boring or draining if too hard. That optimal level just happens to be lower than for ordinary games. More importantly, a drinking game should have a runspace complexity much less than that of ordinary games because impaired memory capacity is one of the first symptoms of intoxication (B). We must avoid a game that is completely not playable (NP-complete).

2.4. Low cost, high reward (PING PONG)

Given the whimsical yet effusive milieu (A) of a party, patrons should not feel too physically, mentally, or emotionally drained after a single game. Therefore we propose the following heuristics to optimally calibrate the primary investment energy gift (or PING) against the principle output and gain (PONG). (1) A single game should not occupy an unreasonably large aliquot of the party time. (2) Players should not have to learn unreasonable skills. (3) Players should receive maximum fun output in exchange for participation input.

Points (1) and (2) requires a reduction in the activation energy for playing the game due to inattentiveness (A) and brainlessness (B). This disqualifies widely lauded games such as Settlers of Catan with the Cities and Knights expansion, Warhammer 40,000, and Dungeons and Dragons. These games may offer high payoff in

the currency of intrigue and imagination, but prove unfeasible for the passing tourist without dedicating hours or weeks preparing and learning the strategy. Unless the social norms of partying undergo a dramatic paradigm shift to accommodate pre-party strategy sessions and avatar development, drinking games will remain limited to simple setup and rules.

To satisfy point (3), players cannot be excessively focused on winning or losing (since only a fraction of players can win each game). Point (3) comes attached with the caveat that anyone who does not find the game “fun” will be ceremonially denominated as “excretory celebrants.” It thus follows that any reasonable partier should find the game entertaining and exciting in a manner linearly related to BAC.

2.5. Drinking is integral (DUI)

We all like games, from corn hole to cricket to Chrono Trigger, and we all like drinking, but drinking games stand alone. While drinking can be performed alongside almost any activity, games that are not designed with drinking in mind often fail to synergize logistically and thematically. Therefore we propose the Drinking is Utterly Indispensable principle (the DUI principle). A drinking game must be unable to progress without players taking their apportioned drinks [1].

For example, while Twister surely makes a fun party game, drinking is at best encouraged but not mandatory. In contrast, flip cup cannot progress until the beverage has been downed (or players start flipping full cups whereupon the game surreptitiously transforms into Stand on a Sticky Wet Floor). Secondly, drinking games are reserved for parties. If one were to play them sober, they would be reduced to “games for people with poor fine motor dexterity” due to (C), or alternatively, “stupidly easy games” due to (B). Third, winning and losing, and increasing inebriation by proxy, should not make the game less fun [1]. In fact, a good drinking game ripens with age as the party progresses!

3. Examples

3.1. Beer Pong

Few drinking games are as popular and time-honored [1] as beer pong, also known as Beirut [10]. Beer pong is often considered the progenitor of the shooting into cups (SIC) drinking games archetype. Thus due to natural selection, one would expect beer pong to be a highly optimized drinking game which satisfies many of our design principles.

With regards to feasibility, beer pong requires virtually no planning (just selecting two or four players), ubiquitous resources (red solo cups and ping pong balls), and little maintenance. However, the full rule set can be quite cumbersome and vary dramatically with the east and west coast populations. Furthermore, the large number of cups poses a high risk of a pathogenic state known as a “party foul.”

Socially speaking, players can freely apply KEG if they can find a substitute, or even take “celebrity shots.” Watching balls land in cups can be as exciting for the players as the crowd. The strategy is simple enough for any patron to enjoy. In fact it may be too simple. We estimate an $O(0)$ runtime complexity for determining the optimal strategy. Beer pong satisfies low-cost, high-reward in many ways. Little preparation and time are necessary. Games can often be decided by the last cup, providing excitement until the last moment.

Drinking is heavily integrated into beer pong, both thematically and mechanically. The cups both hold and are stabilized by the beverage. However this historical methodology has been hotly contested by hygiene scientists. Furthermore, inebriation conveniently amplifies the dexterity challenge. However, one potential issue is that the loser drinks more, becoming less dexterous, which positively feeds back to losing even more.

As we can see, aside from the risk of party foul, excessive simplicity, and potential positive feedback, beer pong is virtually a paragon of design principles. So can we do better? We will demonstrate that improvement is in fact possible.

3.3. Beer Nim

We designed a game entitled beer nim, which is exactly equivalent to the classical game, nim, played with beer cups instead of stones [11]. A number of red solo cups filled with an arbitrary quantity of beverage are arranged into three groups. Players take turns drinking a number of cups (instead of removing a number of stones). The player to drink the last cup wins.

Beer nim requires little planning, simple setup, little maintenance, relatively few and simple rules, and low risk of party foul. However it fails to adhere to, and even actively opposes, virtually all other design principles. Socially, beer nim can only be played with two players, facilitates little conversation because it requires so much thinking, and requires a great deal of attention. It does, however, allow the crowd to vicariously play the game mentally. In terms of difficulty, the runtime complexity of beer nim is technically $O(t)$, though the constant is much larger than the other games discussed. Furthermore, the runspace complexity is significantly larger and left as an exercise to the reader. This is more problematic due to impaired memory constraint.

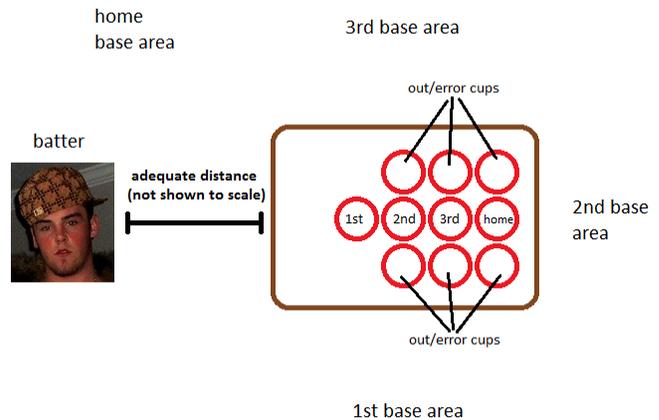
Based on utter failure to satisfy most of the design principles, we must conclude that beer nim is a terrible innovation. Therefore we can use beer nim as a lower bound benchmark.

3.2. Beer Baseball

In preparation for an Olympics themed house party, Gisolfi and Liu-Huang developed a sports-themed drinking game, beer baseball. We describe beer baseball's rules below and compare its funness and adherence to design principles against beer pong, the benchmark.

3.2.1. Setup

- 2 teams of 4+ players (do not have to be the same size)
- Small table
- Line 4 "base" cups moving away from the shooter
- Put 6 additional "out" cups, one on each side of second, third, and home base



3.2.2. Gameplay

- Teams take turns "batting" and "fielding"
- Batting:
 - Players on the batting team take turns trying to shoot the ping pong ball into base cups
 - During her turn, a batter can keep shooting until she makes a base cup or gets out
 - Outs:
 - If the batter misses the cups, he gets a "strike"
 - If a batter gets three strikes, he is out.
 - If the batter ever makes an out cup, he is immediately out regardless of the number of strikes
 - After three outs, the inning ends, and the teams switch batting and fielding roles
 - If the batter makes a base cup, he takes that base by moving to that side of the table (1st base = right side, 2nd = opposite, 3rd = left, home = he goes all the way around)
 - Whenever a batter returns home, each fielder must take a drink
- Fielding:
 - Whenever a batter takes a base, a fielder can choose to make a play
 - If so, she tries to shoot for the same base cup made by the batter
 - If she makes it, the batter is out
 - If she hits an out cup, it is an error, and all the batters advance an extra base
 - If she misses or hits any other cup, nothing happens and she does not get another try

3.2.3. Alternate rules

1. At the start of fielding, each fielder chooses a base and is the only one who can defend that base (requires teams of 4+).
2. If a fielder hits a different base cup than the one made by the batter, it is a “foul.” Nothing happens for a foul; the batter does not get a strike.
3. Whenever the batter misses the cups *but does hit* (anything on) the table, players on the batting and fielding team may both race to retrieve the ball and touch it to the table. If a fielder succeeds, it is a strike. If a batter succeeds, it is a “ball.” If a batter gets two balls, that batter walks to first base for free.

3.2.4. Analysis

Just like beer pong, beer baseball is also a SIC (shooting into cups) game. As such, beer baseball shares the same desirable properties in terms of setup, low-cost high-reward, and integration of drinking. However, beer baseball is more engaging. Players on both the batting team and fielding team have a role to play at all times. Using alternate rule 3, it is even possible to engage all players during each shot. Furthermore, there is nontrivial strategy involved in deciding when to field. Therefore we estimate that the runtime complexity of beer baseball is $O(t)$ with the duration of the game. Having nonzero strategy means the crowd can also engage in discussion and mock strategizing. Considering these points, we believe beer baseball satisfies more design principles than beer pong, and is likely to be a better game.

3.4. Soccer Shots

3.4.1. Setup

- 2 teams (teams must be same size) of 1-3 players (can accommodate even more players, but the table may get crowded)
- Large table
- A ping pong ball
- Two empty six-pack cartons (or some other way to mark the goal)

3.4.2. Gameplay

- Players run around the table using the index and middle fingers of one hand of their choice
- The objective is to flick the ping pong ball into the opposing team’s goal
- Whenever a team scores, the opposing team members must each drink a shot
- No flying: either the index or middle finger must be in contact with the table at all times
- No sliding: you may only move by running along the table using index and middle finger
- Players cannot touch the ball with anything besides the index finger, middle finger, and back of hand of the chosen hand
- If a player breaks a rule, he must drink a shot

3.4.3. Analysis

Among all the games described, soccer shots boasts the easiest setup, requiring only a table, ping pong ball, and two readily available markers (such as a six-pack carton). It is also easy to organize in all other respects, with simple setup and few rules. Socially, soccer inherently requires communication and engages the audience. While soccer shots is easier than soccer, it still requires strategy with respect to formation and coordination. Therefore we estimate that soccer shots has a runtime complexity of $O(t)$. Drinking is not integral because the game is identical without beverage, though “shots” is in the name.

4. Discussion

We sought to codify the core principles common to drinking games. Through close analysis and repeated playthrough of the aforementioned games, we found that the proposed principles are indispensable for a successful drinking game. Through creativity and adherence to the principles, we also designed a drinking game, beer baseball, which satisfies more design principles than even the highly regarded beer pong, our benchmark. While more testing is required, theory suggests that beer baseball is better than beer pong.

5. Acknowledgements

The authors would like to thank Junxing Wang for inspiration and numerous discussions on game design; Nick Gisolfi for helping to compile and prune the list of core principles, and for extensive play testing; and Zachary Batts for extensive play testing.

References

- [1] Brice, M. (25 January 2015). "I went to a drinking game jam and this is what I did." www.mattiebrice.com.
- [2] Horowitz A. (2004). "Dog minds and dog play." In M. Bekoff (Ed.), *Encyclopedia of Animal Behavior*. Greenwood Publishing Group, Westport, CT, 835-838.
- [3] Horowitz, A. (2009). "Domestic dogs (*Canis familiaris*) use visual attention cues when play signaling." *Journal of Veterinary Behavior: Clinical Applications and Research*, 4, 53-54.
- [4] Berghänel, A.; Schülke, O.; Ostner, J. (2015). "Locomotor play drives motor skill acquisition at the expense of growth: A life history trade-off". *Science advances* 1 (7): 1–8. doi: 10.1126/sciadv.1500451
- [5] Robin M Henig (17 February 2008). "Taking Play Seriously". *The New York Times*.
- [6] Fisher-Baum, R. (9 May 2013). "Infographic: Is Your State's Highest-Paid Employee A Coach? (Probably)." *Deadspin*.
- [7] Tom Murphy VII. "New results in k/n Power-Hours." In *Proceedings of SIGBOVIK 2014* (2014).
- [8] g_squidman (2016). "Any Tips for Designing a Good Drinking Game?" *Reddit*.
- [9] "Correlation does not imply causation" *Wikipedia*.
- [10] "Beer Pong." *Wikipedia*.
- [11] "Nim." *Wikipedia*.



CONFIDENTIAL COMMITTEE MATERIALS

SIGBOVIK 2017 Paper Review

Paper 79: Dr. Boozehead, or How I learned to stop worrying and get drunk: Design principles and analysis of drinking games in the silicon age

Robert J. Simmons, second general chair of the 2⁶ish family of SIGsBOVIK

Rating: Unpleasantly sober

Confidence: Mostly shattered by JavaScript

There is yet again a drinking paper that doesn't cite my seminal work. Kids these days.

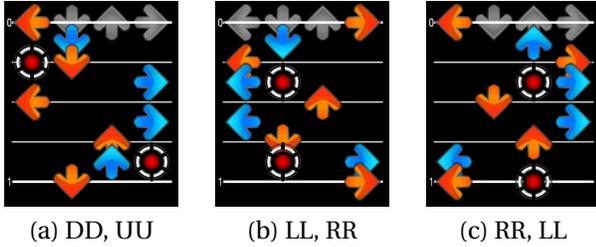


Figure 2. Footswitches of various crossoveriness/facing.

It is tempting to assign the facings L , U , U , and R respectively to the LL , DD , UU , and RR footswitches. However, Figure 2(c) shows that if a footswitch begins with a crossover on U , the facing should be reversed: the RR footing should face L , and LL should face R . “Spin-switches” with D facing are also theoretically possible, arising from patterns such as $LURDDL$ or $LDRUUL$, or similarly, “270-switches”, as shown in Figure 4(a).

Before I realized that, I modified the turniness algorithm (Blum 2016) to face footswitches as above, and it surprised me with charts of $\mathcal{T} > 2$, in excess of the theoretical maximum! I show one such chart in Figure 3(a), in which the step from LL ($\phi = L$) to UL ($\phi = DR$) has individual $\mathcal{T} = 3$, and so on for $UL \rightsquigarrow_R UU \rightsquigarrow_L RU \rightsquigarrow_R RR$. The steps $RR \rightsquigarrow_L LR \rightsquigarrow_R LL$ are both candles ($\mathcal{T} = 2$), resulting overall in $\mathcal{T} = 8/3$ for the whole chart.

Indeed, when I further modified the algorithm to force DD switches to face D (i.e., always facing the direction of the repeated arrow), it produced the chart shown in Figure 3(b), with overall $\mathcal{T} = 3$. (Note its resemblance to the basic spin pattern, $LDRU$, whose $\mathcal{T} = 2$.)

To fairly represent a human player’s desire to step in the least turny of ambiguous ways, I extended the algorithm to provide either the assigned facing, from above, or its polar opposite, chosen at runtime by whichever is closer to the presequent facing. This restores the maximum overall chart turniness to $\mathcal{T} = 2$, a new example of which is shown in Figure 3(c). Figure 4(b) also shows a real-world chart exhibiting this pattern.

However, note that individual steps may still have $\mathcal{T} = 3$, as shown in Figure 3(d). In this example, the step $DL \rightsquigarrow_L LL$ assigns LL to face L , but the subsequent step to LD cannot avoid facing UR . The reason charts still cannot exceed overall $\mathcal{T} = 2$ is that setting up such a situation requires a $\mathcal{T} = 1$ step, which negates the benefit. A chart could conceivably end right before such a step, sneaking through some small ϵ extra turniness (VII 2014) (similar to the case of 270s in (Blum 2016)), but *sustained* average $\mathcal{T} > 2$ remains impossible.

Another approach could assign such a footswitch the opposite footing of the previous facing, regardless of the arrow itself; so in this case the LL would face UL , and each step would have exactly $\mathcal{T} = 2$.

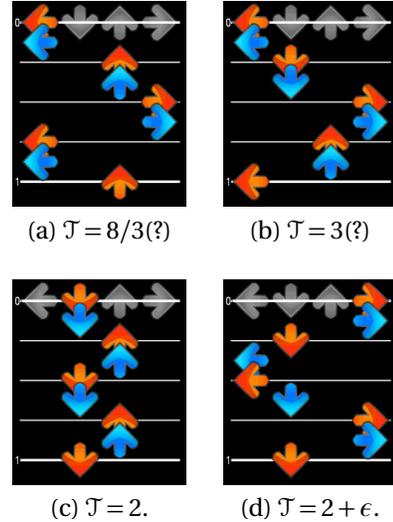


Figure 3. The turniest footswitch patterns. (a) and (b) are false positives (see prose), while (c) and (d) provide theoretically maximal turniness.

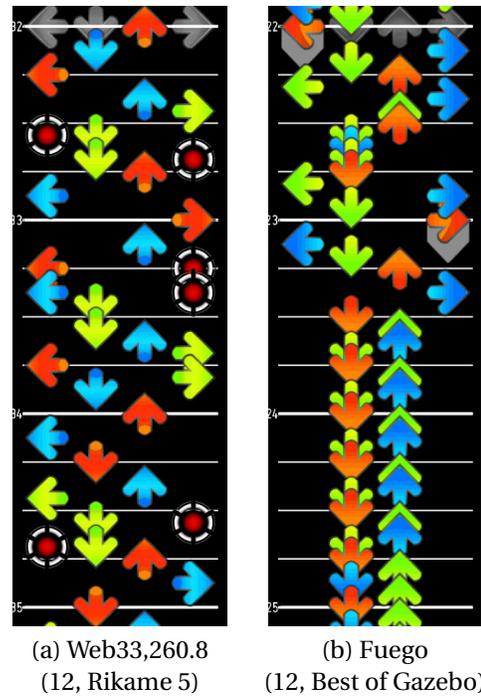


Figure 4. Real-world examples of turny footswitches.

3. Analyzing Crossoveriness

The major flaw of the turniness algorithm (Blum 2016) was that it didn’t care whether a stream started with the left or right foot; it simply exploited the symmetry of Table 1 to find turniness regardless of footing. Hence, it could not distinguish technical footing patterns which

```

data Step = L | D | U | R | Jump deriving Eq

data AnalysisState = S { steps :: Int, xovers :: Int, switches :: Int, jacks :: Int,
    lastStep :: Maybe Step, doubleStep :: Bool, lastFlip :: Bool,
    lastFoot :: Bool, stepsLR :: [Bool] }

commitStream :: AnalysisState -> AnalysisState
commitStream s = s { xovers = xovers s + if f then ns - nx else nx,
    switches = switches s + fromEnum (f == lastFlip s && doubleStep s),
    jacks = jacks s + fromEnum (f /= lastFlip s && doubleStep s),
    lastFlip = f, stepsLR = [] }
    where ns = length $ stepsLR s
          nx = length $ filter not $ stepsLR s
          -- reverse the stream's footing if more L/R steps were crossed over than not.
          f = nx * 2 > ns || nx * 2 == ns && ((switches s > jacks s) == lastFlip s)

analyzeStep :: AnalysisState -> Step -> AnalysisState
analyzeStep s step
    | step == Jump = (commitStream s) { lastStep = Nothing, doubleStep = False }
    | lastStep s == Just step = stream (commitStream s) { doubleStep = True }
    | otherwise = stream s
    where foot = not $ lastFoot s
          -- record whether we stepped on a matching or crossed-over L/R arrow.
          addStep ft L steps = steps ++ [ft]
          addStep ft R steps = steps ++ [not ft]
          addStep ft _ steps = steps -- U/D don't help to determine L/R footing.
          stream s = s { steps = steps s + 1, lastStep = Just step, lastFoot = foot,
            stepsLR = addStep foot step $ stepsLR s }

analyze :: [Step] -> AnalysisState
analyze = commitStream . foldl analyzeStep (S 0 0 0 0 Nothing False False False [])

```

Figure 5. Pseudocode description of the crossoveriness and footswitchiness algorithm.

could affect the way a human would play the chart. It often played charts inhumanly, facing backwards and/or stepping 270s for most of a song.

So, my contribution this year is an algorithm which plays more naturally, and which consequently can report on a chart's technical patterns beyond simple turniness. The algorithm realizes three principles of ITG:

1. Alternate feet as much as possible.
2. Step crossed-over as little as possible.
3. Jumps or jacks allow the player to reset her footing.

Figure 5 describes the algorithm in pseudocode. To summarize it in prose:

- Split the chart into several units of stream, the boundaries of which occur at every jump and any time an arrow is repeated.
- Step each stream with alternating feet.
- Compare the number of matching steps (i.e., *L* foot on *L* arrow or *R* on *R*) versus crossover steps. If the latter

is greater, re-step the stream with opposite feet from before (this kills the crossovers).

- After flipping each stream, if necessary, count the total crossovers in the whole chart.

4. Analyzing Footswitchiness

Because we split the chart whenever an arrow is repeated, figuring out whether that arrow is stepped with different feet on either side of the stream boundary is a natural consequence of figuring out how to step each stream individually. This is also shown in Figure 5's pseudocode. To summarize in prose, if neither stream needed to be flipped (or if both did), then the alternating feet assumption holds, and the repeat must be a footswitch.

5. Analyzing Jackiness

A jack occurs when a repeat arrow is stepped with the same foot, rather than alternating (hence the name, from "jackhammer"). You've got the idea by now, right?

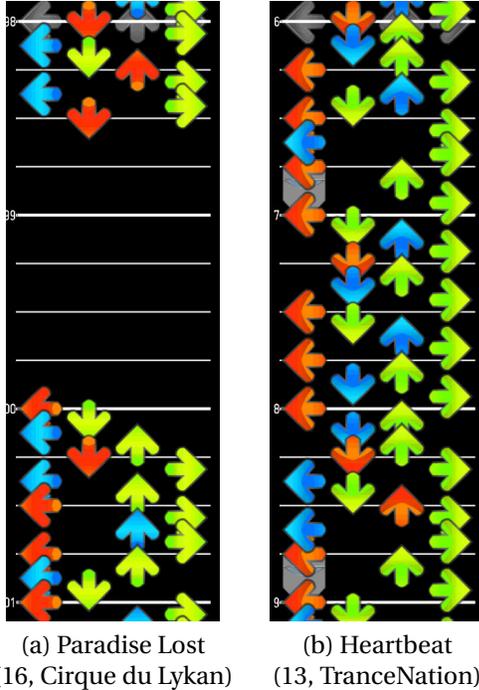


Figure 6. The doublesteps in some streamy charts must be identified, and the stream split, lest “too much” of the following stream appear completely crossed-over.

6. Forced Doublesteps

After painstakingly translating the pseudocode from Figure 5 into a real implementation, I found it vulnerable to false positives when a single double-step could force a long section of stream to be stepped backwards. As an extreme example, consider the pattern $LRLRLRLR-D-LRLRLRLR$. Because no jumps or jacks allow the footing to be reset, either the first or last 4 pairs of LR s must be stepped crossed-over.

Figure 6 shows two examples from real-world charts: in (a), the player must not alternate feet across the measure break, while in (b), two L arrows are replaced with rolls for an artistic visual accent, which must be stepped twice each. On inspection, these charts should be stepped with no crossovers, but were evaluated otherwise (730 XOs (27.9%) and 173 XOs (9.6%), respectively).

To handle such cases, I extended the algorithm with a heuristic to identify when a stream becomes “too crossed-over” for “too long”, and to force a doublestep by splitting the stream to flip it back. Algorithm 1 shows the implementation. I will not summarize how it works due to space limitations, but the description should be intuitive enough. The heuristic evaluated Figure 6’s charts as having 0 crossovers each and 1 and 21 doublesteps, respectively. In my analysis next section, I will use $n_{\text{flip}} = 9$ (determined by inspection of a few favourite charts, which should be scientifically rigorous enough for anyone).

Algorithm 1: HeuristicallyDoublestep(\mathcal{S})

Input : \mathcal{S} , a step sequence $s_0 \dots s_n$
Invariant: $\forall s_i, s_j \in \mathcal{S}, j = i + 1 \rightarrow \neg \text{StreamBoundary}(s_i, s_j)$
Input : n_{flip} , heuristic minimum length
Input : $\%_{\text{flip}}$, heuristic percentage, initially 100

```

1 for  $i \in \text{length}(\mathcal{S}) \wedge \neg \text{defined}(i_{\mathcal{DS}})$  do
2    $S' \leftarrow \{s_k \mid s_k \in \text{LRs}(\{s_j \mid s_j \in \mathcal{S} \wedge j \geq i\}) \wedge k \leq n_{\text{flip}}\}$ 
3   if  $|S'| = n_{\text{flip}} \wedge |\text{Crossovers}(S')| \geq \%_{\text{flip}} \times |S'|$  then
4      $i_{\mathcal{DS}} \leftarrow i$ 
5   end
6 end
7 if  $\text{defined}(i_{\mathcal{DS}})$  then
8   if  $i_{\mathcal{DS}} = 0$  then
9      $i_{\mathcal{DS}} \leftarrow \text{FindUnflippedSection}(\{s_i \mid s_i \in \mathcal{S} \wedge i \neq 0\})$ 
10  end
11  HeuristicallyDoublestep( $\{s_i \mid s_i \in \mathcal{S} \wedge i < i_{\mathcal{DS}}\}$ )
12  HeuristicallyDoublestep( $\{s_i \mid s_i \in \mathcal{S} \wedge i \geq i_{\mathcal{DS}}\}$ )
13 else
14   CommitStream( $\mathcal{S}$ )
15 end

```

7. Evaluation

Our experimental corpus has grown considerably since last year, and now comprises 11,666 stepcharts. I ran the crossoveriness/etcetera algorithm on all of them, and counted the total steps (not including jumps), crossovers, footswitches, jacks, forced doublesteps, and crossover switches for each. I also grouped the charts by author and by song pack to calculate each author’s/pack’s overall crossoveriness/etc. You can view the entire dataset at <http://tinyurl.com/crossoveriest>.

Tables 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, and 12 summarize the dataset as “leaderboards” for each category. They present the data the same way as last year (Blum 2016), so I won’t explain it again. Suffice to say that ITG enthusiasts should use their personal chart style preferences to navigate these tables and find song or pack recommendations or feel proud of themselves or whatever.

In the by-author analysis, I excluded authors with fewer than 10 charts, like last time. Also, in by-author and by-pack, I excluded 1-, 2-, and 3-foot charts, on the pretense that they often ignore the alternating feet assumption (though I also admit this biases the analysis against DDR/Konami). Nevertheless, DDR charts generally ruled the single-digits, even in footswitchiness, showing perhaps more technical depth than I gave them credit for.

By the way, the theoretical maxima for XO%, FS%, and JK% are $50-\epsilon$, $100-\epsilon$, and $100-\epsilon$, respectively (VII 2014).

Ft.	Name	Pack	#XO
6	Autoload	ITG 3	54
7	Tetris	CuoReNeRo M'PacK	58
8	Pulse	CuoReNeRo M'PacK	94
9	MAX Forever	CuoReNeRo M'PacK	107
10	J-PARA SUPER MEGAMIX	CuoReNeRo M'PacK	141
11	Somebody I Used To Know	Best of Gazebo	114
12	Credens Justitiam	Stuff B Likes	136
13	Banshee Strikes	VocaJawnz	152
14	Slow Down	Sexuality Violation 2	160
15	yoshikawa45 vs siesta45	Rikame's Simfiles 4	111
16	Your Best Nightmare	Undertale	97
<i>no 17s-19s with ≥ 50 XOs</i>			
20	Rainbow Dimension	Rikame's Simfiles 2	84
21	Teenage Dream	Sexuality Violation 2	280

Table 2. Charts with the most total crossovers (XOs).

Ft.	Name	Pack	XO%
4	DAM DIRIRAM	DDR 3rdMIX	27.3
5	STRICTLY BUSINESS	DDR 1st	21.4
6	STRICTLY BUSINESS	DDR 1st	20.9
7	MOBO*MOGA	DDR EXTREME	17.3
8	PARANOiA	DDR 1st	19.6
9	Dazzlin Darlin	r21twins	22.0
10	Enchanted Journey	ITG Rebirth	19.3
11	Lune Noir	r21freak Friendship	13.4
12	W'peg is Fucking Over	best of r21freak ii	14.3
13	The Sampling Paradise	The Paradise Sampler	15.6
14	Slow Down	Sexuality Violation 2	13.3
15	yoshikawa45 vs siesta45	Rikame's Simfiles 4	10.8
<i>no 16s-19s with $\geq 10\%$ XO</i>			
20	Rainbow Dimension	Rikame's Simfiles 2	10.7
21	Teenage Dream	Sexuality Violation 2	11.8

Table 3. Charts with the highest percentage of XOs among total steps.

8. Discussion

To verify the algorithm's accuracy, I manually inspected a random (read: not random) sample of the charts at or near the top of the various leaderboards (read: I played a lot of Stepmania). I also consulted a leading expert in the field of automated ITG chart analysis (read: myself), who reported that the algorithm is infinitely more accurate than the prior state-of-the-art.

Honestly though, it works really well. Last year's algorithm was often finicky and prone to all sorts of false-positives, while this one plays ITG in a recognizably human way (almost) without fail. It was a joy to use.

Surprises. On occasion, the algorithm surprised me by stepping with crossover footswitches which, at first glance, I would probably jack or double-step. However, these always proved to be perfectly valid alternative footings, in some cases requiring considerable look-ahead.

Ft.	Name	Pack	#FS
6	Sweat Shop	ITG Rebirth 2 Beta	56
7	Silent Hill	DDR 3rdMIX	29
8	Pulse	CuoReNeRo M'PacK	41
9	Dr. Boom-Bombay	fort rapids vii	34
10	Eat 'Em Up!	Mute Sims 5	54
11	Nemeton	The Legend of Zim 4	97
12	Nemeton	Subluminal	140
13	Love Is Eternity	Subluminal	140
14	Switch	Getty	347
15	Danse Macabre	Aoreo's Ariginals	201
16	Weird Science	Stamina Showcase	61
17	Arcane Apparatus	Tachyon Gamma	32
18	Metallic-A-	Oh Henry! Mad Stamina	27
19	Geronimo	Sexuality Violation 3	39
20	Scatman's World	Jimmy Jawns 2	22
21	He He He	Jimmy Jawns 2	4
22	Architecture	SPEEEDCOOOORE 4	24
23	Geronimo	Sexuality Violation 3	39

Table 4. Charts with the most total footswitches (FSs).

Ft.	Name	Pack	FS%
3	Sweat Shop	ITG Rebirth 2 Beta	15.3
4	DROP THE BOMB	DDR 3rdMIX	9.8
5	MAKE IT BETTER	DDR 2ndMIX	20.0
6	Sweat Shop	ITG Rebirth 2 Beta	18.2
7	5.1.1	DDR MAX	12.9
8	La Señorita Virtual	DDR 3rdMIX	8.4
9	PARANOiA KCET	DDR 2ndMIX	10.2
10	Delhi Ill	Mute Sims 6	11.6
11	Sweat Shop	ITG Rebirth 2 Beta	11.8
12	Nemeton	Subluminal	14.8
13	Love Is Eternity	Subluminal	16.8
14	Switch	Getty	15.7
15	Flames of the Sky	fort rapids vii	16.5
16	Mermaid Island	Tachyon Alpha	9.5

no 17s+ with $\geq 5\%$ FS

Table 5. Charts with the highest percentage of FSs.

Ft.	Name	Pack	#XF
<i>no 8s- with ≥ 12 XFs</i>			
9	Dr. Boom-Bombay	fort rapids vii	18
10	Toxic	Sexuality Violation 2	12
11	Heart Shooter	VocaJawnz	44
12	Web 33,260.8	Rikame's Simfiles 5	16
13	Toxic	Sexuality Violation 2	12
14	Fancy Footwork	Cirque du Zeppelin	40
15	yoshikawa45 vs siesta45	Rikame's Simfiles 4	20

no 15s+ with ≥ 12 XFs

Table 6. Charts with the most crossover footswitches (XFs). (Here I chose 12 as the cut-off to exclude a bunch of ambiguously-patterned charts from DDR.)

Author	Charts	Total Steps	XO%
Konami	530	114623	8.03
sssmsm	41	17219	6.55
NEMORIGNAL	44	20165	5.36
M. Emirzian	23	9307	5.16
J. DeGarmo	16	4362	4.86
D. Renzetti	18	7877	4.47
R. McKanna	47	14855	4.23
M. Puls	26	8379	4.18
King of Light	24	8817	4.13
D. Bernardone	217	76290	4.10
D. D'Amato	107	36265	3.92
bblum	32	32382	3.76
...			
B. Vergara	13	15454	0.091
Aoreo	21	27990	0.089
Zaia	368	448460	0.080
t0ni	85	128964	0.058
Burn	27	60052	0.057
Dirk	12	21996	0.055
Happy Feet	30	57888	0.040
@@	63	199530	0.026
Arvin	79	108612	0.023
Drazu	153	221460	0.021
teejusb	11	11298	0.018
Fraxtil	19	25612	0

Table 7. Chart authors with the highest/lowest XO%.

Author	Charts	Total Steps	JK%
King of Light	24	8817	12.94
R. McKanna	47	14855	9.95
Konami	530	114623	9.63
P. Shanklin	21	9834	8.91
M. Puls	26	8379	8.90
K. Ward	281	86475	8.80
ATB	31	20673	8.46
J. DeGarmo	16	4362	8.30
D. Bernardone	217	76290	7.98
Renard	45	15035	7.94
C. Foy	133	52418	7.72
Yoko	10	4128	7.63
...			
bblum	32	32382	1.56
B. Vergara	13	15454	1.31
Arvin	79	108612	1.30
Drazu	153	221460	1.06
Zaia	368	448460	0.86
Aoreo	21	27990	0.78
T. Swag	13	22909	0.70
@@	63	199530	0.61
warpdrive	16	19090	0.49
Dirk	12	21996	0.35
Burn	27	60052	0.28
Hsarus	18	70162	0.27

Table 9. Chart authors with the highest/lowest JK%.

Author	Charts	Total Steps	FS%
Konami	530	114623	2.29
bblum	32	32382	1.29
M. Puls	26	8379	1.16
R. McKanna	47	14855	1.11
mudkyp	63	42792	1.09
S. Venkat	24	11084	0.96
K. Ward	281	86475	0.89
xRGTmX	19	11734	0.88
sssmsm	41	17219	0.85
D. Bernardone	217	76290	0.83
ATB	31	20673	0.82
Happy Feet	30	57888	0.82
...			
Hsarus	18	70162	0.070
Drazu	153	221460	0.057
@@	63	199530	0.037
Revolver	11	8302	0.036
T. Swag	13	22909	0.031
Dirk	12	21996	0.023
B. Vergara	13	15454	0.019
warpdrive	16	19090	0.016
Burn	27	60052	0.015
teejusb	11	11298	0.009
t0ni	85	128964	0.002
S. Tofu	26	34609	0

Table 8. Chart authors with the highest/lowest FS%.

Pack	Charts	Total Steps	XO%
DDR 1st Mix to Extreme	530	114623	8.03
r2112	47	18377	4.50
the best of r21freak	100	45156	4.22
the best of r21freak ii	48	25024	4.16
In The Groove 2	222	66113	4.05
In The Groove Rebirth+	108	43758	3.99
r21twins	52	22347	3.89
In The Groove 3	320	106363	3.61
CuoReNeRo MeGaPacK	423	248625	3.45
In The Groove	1408	491819	3.27
r21freak Friendship Pack	47	20363	3.13
BemaniBeats 4	31	18464	3.02
...			
Tachyon Epsilon	150	208830	0.064
SPEEEDCOOOORE 4	101	123814	0.064
TranceMania	80	121415	0.062
Cirque du Lykan	129	160312	0.059
Cirque du Zonda	45	74890	0.057
Jimmy Jawns	109	170894	0.044
Getty	26	53528	0.043
Tachyon Delta	32	36712	0.038
Tachyon Gamma	32	36134	0.033
Oh Henry! Mad Stamina	46	152326	0.028
Causality Violation	10	19507	0.021
Fast Track to Brutetown	29	46082	0.020

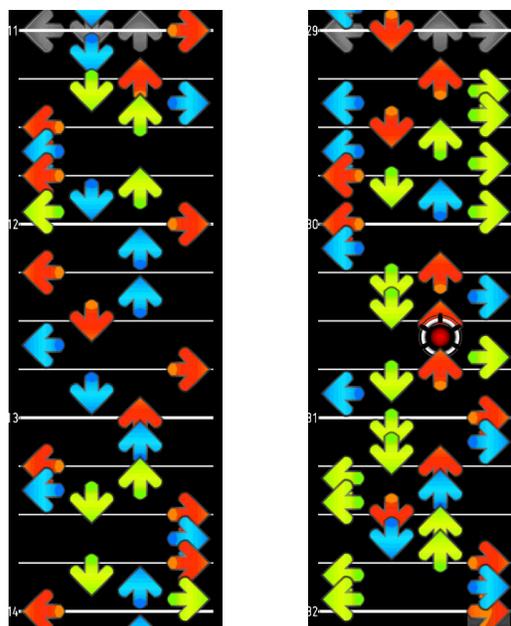
Table 10. Packs with the highest/lowest XO%.

Pack	Charts	Total Steps	FS%
Subluminal	17	13418	4.70
Aoreo's Ariginals 2	16	15222	2.42
DDR 1st Mix to Extreme	530	114623	2.29
Aoreo's Ariginals	31	26418	2.26
rocky mount xi	113	79986	0.97
In The Groove 2	222	66113	0.93
FA and Chill	35	22045	0.86
Getty	26	53528	0.85
Undertale	19	14722	0.84
r2112	47	18377	0.82
Fort Rapids VI	75	75563	0.77
Mute Sims 8	72	47140	0.71
...			
SPEEDCOOOORE 4	101	123814	0.093
Stamina Showcase	38	126146	0.088
VocaJawnz II	128	183153	0.084
Cirque du Zeppelin	109	102991	0.082
SPEEDCOOOORE 3	66	69236	0.071
Causality Violation	10	19507	0.051
TranceNation	41	123940	0.047
Oh Henry! Mad Stamina	46	152326	0.046
Tachyon Epsilon	150	208830	0.040
Noisiastreamz	20	41917	0.019
TranceMania 2	40	64109	0.003
TranceMania	80	121415	0.001

Table 11. Packs with the highest/lowest FS%.

Pack	Charts	Total Steps	JK%
DDR 1st Mix to Extreme	530	114623	9.63
r2112	47	18377	8.80
In The Groove 2	222	66113	8.64
Gensokyo Holiday	87	51652	7.87
r21Freak's Friendship Pack 2	32	15649	7.71
Omnifarious	10	5594	7.53
r21twins	52	22347	7.18
Piece of Cake 7	20	11554	6.97
In The Groove	1408	491819	6.97
In The Groove Rebirth+	108	43758	6.97
TLOES Chapter 1	85	42201	6.89
ITG Rebirth 2 Beta	262	99078	6.79
...			
TranceMania 2	40	64109	1.03
Causality Violation	10	19507	0.98
VocaJawnz II	128	183153	0.98
Tachyon Epsilon	150	208830	0.94
Tachyon Delta	32	36712	0.87
Cirque du Lykan	129	160312	0.87
Cirque du Veyron	31	51545	0.79
Cirque du Zeppelin	109	102991	0.77
Oh Henry! Mad Stamina	46	152326	0.67
Stamina Showcase	38	126146	0.56
Cirque du Zonda	45	74890	0.47
TranceNation	41	123940	0.25

Table 12. Packs with the highest/lowest JK%.



(a) Dr. Boom-Bombay (9, fort rapids vii) (b) Toxic (10/13, Sex'y Violation 2)

Figure 7. Sometimes the algorithm was smarter than me.

In Figure 7(a), the chart repeats *L* (later *R*) thrice, beginning with the right (later, left) foot. While a human player would jack these repeated arrows, the crossoveriness algorithm performs a double-footswitch, effectively reducing the total crossover steps by 1 each time. In Figure 7(b), a mine cues the player to double-step with her right foot, but the crossoveriness algorithm can begin this section already crossed-over, owing to an earlier *L* jack on which it could switch feet.

Honourable Mentions. I omitted a table for the jackiest charts, on account of most of them being either trivial beginner charts or extra-long megamixes. One deserves a special mention: Sandstorm (Jimmy Jawnz 2), shown in Figure 8(a), has more than twice as many total jacks as the next jackiest chart, clocking in at 1049 (78.5%) with its 15 and 992 (69.6%) with its 17. And looking at that chart, can't you just hear Sandstorm playing in your head already?

I also wanted to highlight the chart with the most crossover switches, shown in Figure 8(b), mostly because the skittle notes should add some nice variety of colour to the paper (with apologies to the dead-tree SIGBOVIK audience reading in greyscale). Figure 8(c), with 2nd place in crossover switches following (b), comes with an edit chart titled "no sidefoots", and to be perfectly honest I just kept saying the word "sidefoots" to myself and giggling a lot while writing this paper.

Sweet spot. Finally, in case it wasn't obvious in the tables, I'll point out that 9-15 is clearly the sweet spot of difficulties for technical stepcharts.

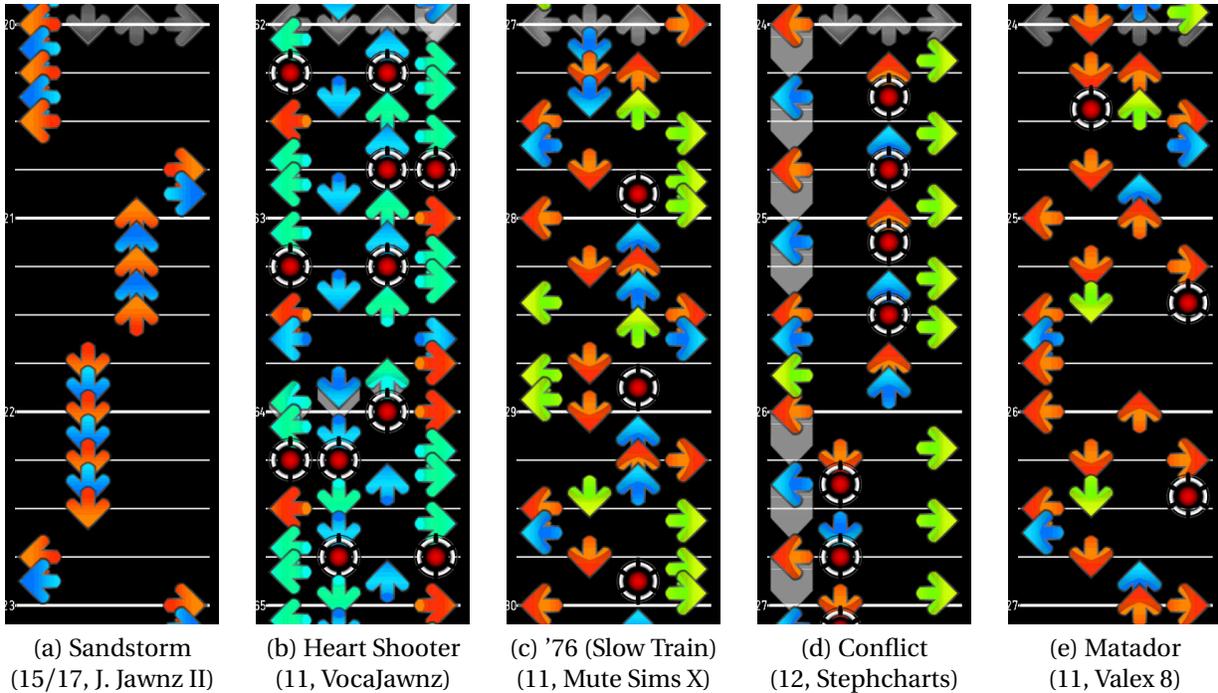


Figure 8. Miscellaneous interesting charts I discovered while browsing the giant spreadsheet.

9. Never Work

Let’s be honest: this isn’t gonna be a paper trilogy. Okay, with that said, here are some things that would be cool to implement in a fantasy universe with infinite free time. (I have renamed this “future” work section accordingly.)

There are a few remaining cases the algorithm doesn’t yet understand:

- Doublesteps forced either by mine cues or by holds;
- Crossover and/or bracket jumps, not usually forced but often way less turny than the alternative;
- Forced footings across stream boundaries arising from bracket jumps or jump-footswitches.

For example, Figure 8(d) shows many sequential doublesteps, each forced by a mine cue, but which the algorithm interprets as spins because the flipped stream length falls below n_{flip} . Figure 8(e) shows an example of jump-footswitches which the algorithm fails to count because it ignores the footing of jumps.

These patterns would all have to be identified heuristically. Apart from that being more work than I wanted to do, I also feel that adding too many heuristics to SIGBOVIK research compromises the simple and innocent beauty of an implementation unbound by the demands of mainstream conferences.

10. Conclusion

Please accept my paper. I worked hard on it.

References

- B. Blum. Which ITG stepcharts are turniest? SIGBOVIK, 2016.
- T. VII. What, if anything, is epsilon? SIGBOVIK, 2014.



Dog track

Nonstandard ML



- 13 **Batch normalization for improved DNN performance, my ass**
Joshua A. Wise

- 14 **Colonel density estimation**
Harish Krishna *et al.*

- 15 **Degenerative adversarial networks**
Raphael Gontijo Lopes and Diptodip Deb

- 16 **Stopping GAN violence: Generative unadversarial networks**
Samuel Albanie, Sébastien Ehrhardt, and João Henriques

Batch Normalization for Improved DNN Performance, My Ass

Joshua A. Wise
joshua@joshuawise.com
Emarhavi Heavy Industries

Abstract

Batch normalization is an extremely popular technique to enable faster training, and higher network performance after training. We apply batch normalization to a relatively small network, and find it to be completely ineffective, and indeed, to reduce network convergence and overall network performance.

1. Introduction

Batch normalization [4] is a strategy used to accelerate learning in deep neural networks. Theorized to work by reducing “internal covariate shift”, it dynamically computes normalization coefficients at each channel internal to a convolutional network while training, and then during validation and operation, hopes that they generalize. Although the effect of batch normalization can, in theory, be baked into weights at each neuron, the batch normalization coefficients are not learned through gradient descent, and only their second-order effects are. Through a convoluted process, this means that adding more parameters somehow makes the network converge more readily, and so everybody does it.

Batch normalization has been used in many networks from deep to shallow: recent DCGAN architectures (for instance, `pix2pix` [5]) have used batch normalization between layers when training regression, and Google’s Inception network has used it when training classification. Batch normalization is said to be tolerant to hyperparameters; for instance, the `decay` hyperparameter is said to reasonably range from 0.999 through 0.99 all the way down to 0.9 and “etc.”, which is apparently one nine fewer than 0.9. It also has a configurable value of `epsilon` [2], which is likely to be valuable during times of shortage [9].

In this work, we sprinkle batch normalization pixie dust onto an existing neural network to improve its performance, and analyze the performance gained.

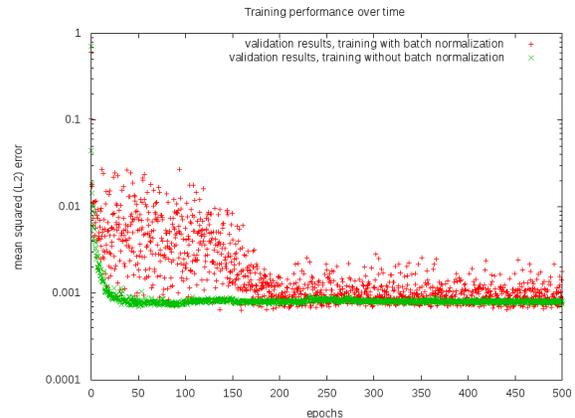


Figure 1. Batch normalization performance vs. classical training.

2. Related Work

Everybody who does work on deep neural networks cites the founding paper on the subject that was written long before anyone had ever heard of a GPU. So we do so here too [7]. But let’s be real here, this whole lab report is actually a take-off of Kovar and Hall [6], who did this way better than I did.

3. Experimental Procedure

We took an existing neural network of a few layers, a corruption of the work in [3]. It already did not work very well, but the batch normalization pixie dust was expected to substantially improve it, and make everything all better. We inserted batch normalization layers in all but the final convolutional layer, since adding a normalizing layer before the output seemed obviously stupid and likely to produce absurd nonlinearities.

The batch normalization layer was built using TensorFlow’s `tf.contrib.layers.batch_norm` [10] function. (The `contrib` in the Python module path means that the routine is extra-well-tested.) We experimented with multiple sets of hyperparameters, primarily because the first set of hyperparameters were no good. The initial set of hyperparameters used a value of 0.9 for `decay` and a value of 10^{-5} for `epsilon`, because

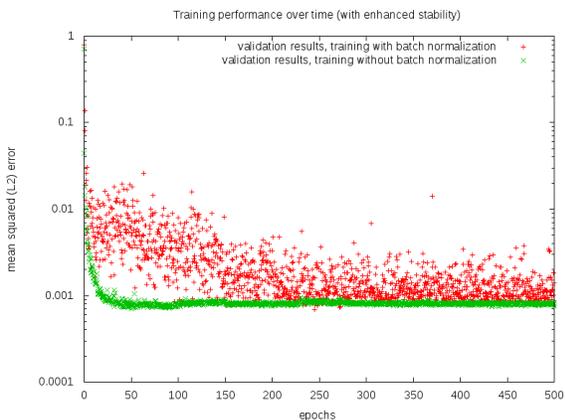


Figure 2. Batch normalization performance, with stability enhancements.

that’s what pix2pix did. The results were, hey, wait, this is the wrong section for that.

The second set of hyperparameters used increased the decay coefficient to 0.999, and enabled `zero_debias_moving_mean`, because it is said that one should do that if one’s results are unstable.

Training on both runs took place overnight using TensorFlow on 8 NVIDIA Really Big GPUs in parallel. On the new power-efficient Pascal architecture, training consumed approximately 1.5 kW, for 12 hours, or 18 kWh of total power, or enough for my coworker to boil 540 cups of tea [1].

4. Results

The results were utter crap. The first run was dramatically unstable (see Figure 1). When measures were taken to make the system more stable, it responded in the opposite fashion (see Figure 2). Convergence did not happen faster than without batch normalization, inasmuch as anything that the batch normalization runs did could be at all described as converging.

Visual quality of the output batch-normalized runs was not verified, because, let’s face it, it’s going to be noisy crap. Also, I didn’t finish writing the support to load and save the batch-normalization coefficients into checkpoint files, so that’s another strike against that.

5. Future Work

Maybe someone can get this crap to work. Like, everyone else who sprinkles batch normalization pixie dust on their CNNs gets them to train right quick, and the Google folks say that you don’t even need L_2 normalization with them, let alone any other kind of normalization. Work should be done to investigate whether the Google folks just got a really lucky RNG seed each time they did their batch-norm runs, or maybe a really bad one for their control runs, because clearly this stuff ain’t working for me.

Other experiments could be run with other normalization schemes, like Dropout [8]. Initial experiments are under way that indicate that all of the literature about Dropout is also a lie.

6. Conclusion

I still don’t know anything about how neural networks work, and as far as I can tell, neither does anyone else.

References

- [1] Personal correspondence.
- [2] Tom 7. What, if anything, is epsilon? *SIGBOVIK*, you know, like the only one that year, 2014.
- [3] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image super-resolution using deep convolutional networks. *CoRR*, abs/1501.00092, 2015.
- [4] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *CoRR*, abs/1502.03167, 2015.
- [5] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-image translation with conditional adversarial networks. *CoRR*, abs/1611.07004, 2016.
- [6] Lucas Kovar. Electron band structure in germanium, my ass. Online, <http://pages.cs.wisc.edu/~kovar/hall.html>, 2007.
- [7] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel. Backpropagation applied to handwritten zip code recognition. *Neural Comput.*, 1(4):541–551, December 1989.
- [8] Infected Mushroom. Drop out. From the album, *Converting Vegetarians, Disc 2*, 2003.
- [9] Chris Tuffley. The great epsilon shortage. *Mathematical Intelligencer*, 21(4):37, 1999.
- [10] Someone who didn’t proofread their code sample. Tensorflow API documentation. Online, https://www.tensorflow.org/api_docs/python/tf/contrib/layers/batch_norm, 2017.

Colonel Density Estimation

Harish Krishna
harishkrishna.v@research.iiit.ac.in
Bharat Lal Bhatnagar
bharat.bhatnagar@research.iiit.ac.in
Nishant Prateek
nishant.prateek@research.iiit.ac.in
Bhaktipriya Shridhar
bhaktipriya.r@students.iiit.ac.in

Kohli Center for Intelligent Systems
IIIT-H
Hyderabad

Abstract

The highly relevant and important problem of Colonel Density Estimation has seen little focus in recent times. In this work, we present fresh approaches to solve both classes of Colonel Density Estimation - Colonel-Density Estimation and Colonel Density-Estimation. The proposed solution is currently the state-of-the art in both classes of the problem. We also discuss how this approach can easily be extended to solve the more General Density Estimation problem.

1 Introduction

Colonel Density Estimation is an important problem in the fields of statistics, physics, biology and military recruitment but has surprisingly got almost no interest from research groups. In contrast, the easier and similar sounding problem of Kernel Density Estimation has seen a significantly disproportionate amount of effort trying to advance the current state-of-art. Recently, [4] suggested the method of diffusion while [12] explained how to go about choosing the kernel and bandwidth.



Figure 1: A colonel

Despite the apparent similarities of the two problems, the methods used for Kernel Density Estimation can't be used for Colonel Density Estimation [11]. Kernel Density Estimation can give only the density of the bones of the colonel, not the whole colonel. Finding the density of a whole colonel is one of the two classes of problems Colonel Density Estimation must solve. Also, while Kernel Density Estimation is a non-parametric method, Colonel Density Estimation is non-paramilitaric. The

method suggested in this paper is more akin to generalization when compared to the existing approaches for Kernel Density Estimation.

There are two classes of Colonel Density Estimation. The first, *Colonel-Density Estimation* is the problem of estimating the density of colonels. The second, *Colonel Density-Estimation* involves using colonels for the problem of Density Estimation. In this work, we provide a novel approach for Colonel-Density Estimation that beats all prior research that attempts to solve this problem. For the second problem, we propose a solution that needs much fewer resources when compared to the state-of-the-art.

2 Colonel-Density Estimation

We solve the problem of estimation of the density of colonels by first finding the mass and volume of the colonel and dividing the two quantities to get the density.

2.1 Finding mass

We were disappointed that though several earlier works like [7], [9], [3] claim to introduce a novel method, they do not actually use any novels. We introduce a novel method that does actually involve novels. Though we are the first to use novels for mass in the context of density estimation, the idea we propose and the novel have been time-tested in a different field for several centuries now.

Inspired by religion where a novel is used for obtaining mass, we let out colonels read out from the same novel. The money raised in the process (but expressed in the SI units of mass) is the mass of the colonel.

2.2 Finding volume

We use the classical method [2] to find the volume of a colonel. The colonel is immersed into a tank filled to the brim with a Newtonian liquid. The volume of the colonel is equal to the volume of the liquid displaced. We experimentally found that better results were achieved when the colonel was immersed for quite a while so that the liquid displaces the air in the lungs of the colonel as well.

2.3 Calculation of density

We calculate density as

$$D = \frac{\text{mass}}{\text{volume}} \quad (1)$$

where the mass and volume are in SI units (when the volume is zero, it would mean that Nishant had probably messed up somewhere.)

2.4 Results

We choose colonels who know their densities for evaluating our performance. The calculated density D of a colonel is correct if it falls between $\hat{D} - \epsilon$ and $\hat{D} + \epsilon$ where \hat{D} is the actual density of the colonel. The accuracy of the method is equal to the ratio of the number of correct estimates of density to the total number of colonels who participated in the experiment.

We compare our performance with [8], [1] and [6]. The results are summarized in Figure 2. It is to be observed that we perform significantly better than other methods. We reason that this could be because these earlier works did not intend to solve the problem of Colonel-Density Estimation at all. We leave this for future work to verify.

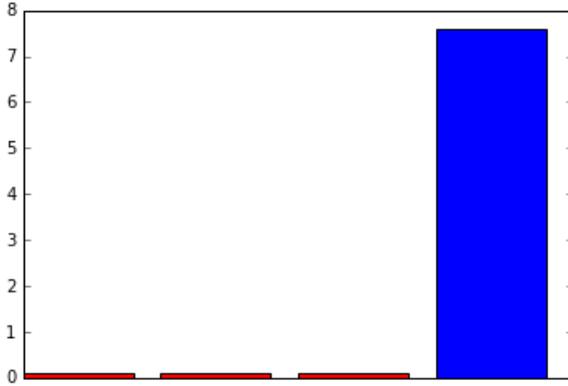


Figure 2: Comparison of accuracies of our method (blue) with respect to other methods (red) for the task of Colonel-Density Estimation

3 Colonel Density-Estimation

There is very limited prior knowledge and experiences in using colonels for density estimation. Colonels, usually found shouting for attention, are easy and efficient tools for density estimation. In our experiments, we found that if we asked a colonel politely to guess the density of an object we were pointing to, the colonel usually obliged. However, evaluation of colonel-based density estimation is hard [5] as the accuracy of the method is so heavily dependent on the choice of colonel and how annoyed the colonel is. This is one similarity this shares with Kernel Density Estimation where the choice of Kernel makes a difference.

Instead, we assert the relevance of our approach by comparing the kind of resources our method needs with prior work that uses colonels for density estimation. The only works that we found that could perhaps give a more accurate estimate of density using a colonel are in [13] and [14]. We believe that the method of just asking the War Machine with Colonel J Rhodes in it has the potential to be more accurate than our method. This is because Jarvis, the Artificial Intelligence that helps control the metal suit, is pretty smart and probably knows the densities of most objects. However, the cost of building such a War Machine or Iron Patriot suit is very high [10] and can only be afforded by billionaires. In comparison, the cost of our suggested method is negligible (refer Figure 4). Also, since Colonel J Rhodes is currently recovering after an injury sustained in a civil war, our colonel-based density estimation technique is the state-of-art, at least until he returns.



Figure 3: Colonel Rhodes and Jarvis, the only competition for Colonel Density-Estimation

4 Generalizability

The solutions to the problem of Colonel Density Estimation proposed in this paper are perhaps among the most easily generalizable solutions ever. This only involves replacing the colonel with a general in every step of each process. We found that in general, generals perform better at density estimation. An interesting observation was that the colonel-density and general-density are different, despite both colonels and generals being humans. We attribute this to the fact that generals are more mean, and hence probably more thick-skinned.



Figure 4: Colonel Rhodes in his War Machine suit was the state-of-art for Colonel Density-Estimation until he was injured.

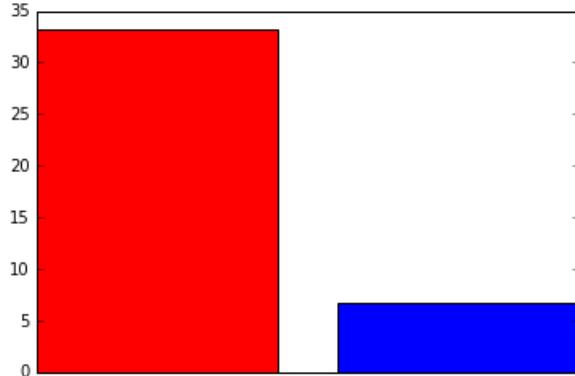


Figure 5: Cost of [14] (red) and our proposed solution (blue) for Colonel-Density Estimation. Note that the y-axis is a logarithmic scale.

5 References

- [1] Larry C Andrews and Ronald L Phillips. *Laser beam propagation through random media*, volume 1. SPIE press Bellingham, WA, 2005.
- [2] Archimedes. Eureka eureka, 220BCE.
- [3] Geoffrey H Ball and David J Hall. Isodata, a novel method of data analysis and pattern classification. Technical report, DTIC Document, 1965.
- [4] Zdravko I Botev, Joseph F Grotowski, Dirk P Kroese, et al. Kernel density estimation via diffusion. *The Annals of Statistics*, 38(5): 2916–2957, 2010.
- [5] Theophilos Cacoullos. Estimation of a multivariate density. *Annals of the Institute of Statistical Mathematics*, 18(1):179–189, 1966.
- [6] Ian Holyer. The np-completeness of edge-coloring. *SIAM Journal on computing*, 10(4):718–720, 1981.
- [7] Kazutaka Katoh, Kazuharu Misawa, Kei-ichi Kuma, and Takashi Miyata. Mafft: a novel method for rapid multiple sequence alignment based on fast fourier transform. *Nucleic acids research*, 30(14):3059–3066, 2002.
- [8] Rainer Martin. Noise power spectral density estimation based on optimal smoothing and minimum statistics. *IEEE Transactions on speech and audio processing*, 9(5):504–512, 2001.
- [9] Nicholas J Miller, Catherine Rice-Evans, Michael J Davies, Vimala Gopinathan, and Anthony Milner. A novel method for measuring antioxidant capacity. *Clinical science (London, England: 1979)*, 84(4):407–412, 1993.
- [10] moneysupermarket.com. The cost of building one iron man suit.
- [11] Nishant Prateek. On the differences between colonel density estimation and kernel density estimation, 2006.
- [12] Simon J Sheather and Michael C Jones. A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society.*, pages 683–690, 1991.
- [13] Marvel Studios. Iron man movies, 2007-2017.
- [14] Marvel Studios. Avengers movies, 2011-2017.

Degenerative Adversarial Networks

Raphael Gontijo Lopes* & Diptodip Deb †
Georgia Tech, Atlanta, GA
{raphaelgontijolopes, diptodipdeb}@gatech.edu

Abstract

In recent years, Deep Learning researchers have collectively achieved a pace of useful information extraction that is dangerously close to outstripping the second law of thermodynamics. To solve this problem, we propose a new framework for estimating degenerative models via an adversarial process, in which we simultaneously train two models: a degenerative network D that destroys the data distribution and a discriminative model D that estimates the probability that a sample came from true noise rather than D . The training procedure for D is to maximize the probability of D making a mistake. Within the space of arbitrary D and D , we roll a D20 and check for damage. This system corresponds to entropy maximization, which ensures a timely heat death. Experiments would have demonstrated the potential of the framework, but most of our results were degenerated in the process of running them.

1 Introduction

The promise³ of deep learning is to discover rich, hierarchical models [5] that represent probability distributions over different kinds of data, such as natural images, audio waveforms containing speech, and symbols in natural language corpora (see Figure 1). All of this structuring of data works to decrease entropy by creating discriminators that are able to classify this data into well-defined labels. Furthermore, we now see the success of deep generative models due to Goodfellow et. al [5] and Kingma et al. [8], which further accelerates the pace of structured data generation.



Figure 1: The power of deep learning. [photos: Bobolas, 2009 [1], Maley, 2011 [10]]

All of this model discovery is creating too much information. At this rate, we will outstrip the second law of thermodynamics and begin to decrease entropy in the universe (see Figure 2). In order to maintain reality as we know it, we present the Degenerative Adversarial Network, or DAN, which sidesteps the successes of deep learning models in order to maintain a steady and healthy pace towards the sweet release of heat death.

In this proposed adversarial degeneration framework, the degenerative model is pitted against an adversary: a discriminative model which attempts to distinguish whether actual data has been degenerated or whether the observed sample is true noise. The degenerative model can be thought of as a team of steamrollers, flattening data into a uniform distribution for maximum entropy. The discriminative model can be thought of as a team of protractors, trying to determine if its input has been properly flattened into true noise. Competition in this game will drive both groups to improve until the discriminator cannot reliably distinguish between generated noise and degenerated data.

*Currently looking for grad school.

†Please accept me into your lab.

³unfulfilled

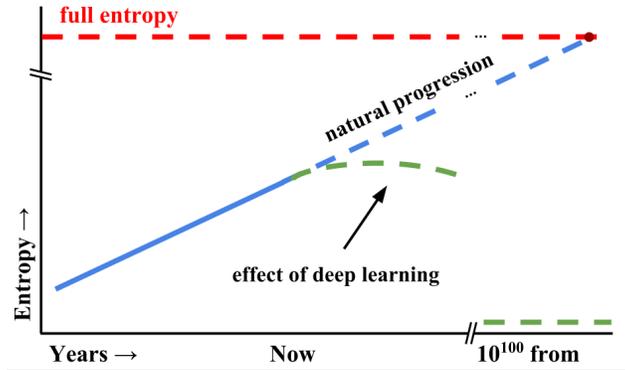


Figure 2: Plot showing the dangers of deep learning. In hindsight, this plot generates information so please refrain from looking at it.

This framework can give specific algorithms for degeneration and discrimination. We explore the case in which the degenerative model destroys data by passing it through a multilayer after being perturbed by noise and the discriminator model is also a multilayer perceptron. We refer to this special case as a *DAN* and show that we can train both networks using backpropagation in an end-to-end fashion. Uniquely to our approach, there is no need to actually code this network.

2 Related Work

Training adversarial networks is famously hard [12], because the optimization objective equates to trying to find a Nash equilibrium in a non-cooperative game. We found this to be even more complicated when degenerating, because the procedure makes it very easy for the Degenerator to output images of PhD students⁴.

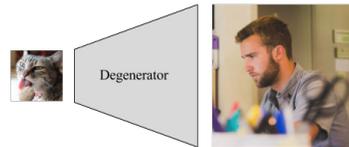


Figure 3: Example of bad Nash Equilibrium for the DAN model, the degenerate result is a PhD candidate hard at "work" on his thesis.

To solve this, some work in the field has argued that a balance between the two adversaries needs to be found in order to stabilize the game and avoid local minima. However, Goodfellow [4] shows how a more robust solution involves creating an overpowered adversary Discriminator, which ensures that the values of collaboration and looking past one another's differences will not come into play. Therefore, we proceed with the adversarial technique. Our methods are described below.

Goodfellow et al. [6] also use adversarial techniques to degenerate data. However, their method is limited in that it is only able to trick other neural networks. While this is useful for slowing down the pace of data creation by generative models, it is not sufficient for our objectives as their results preserve enough structure that humans can still discriminate with ease. The method we present is robust even to human discriminators.

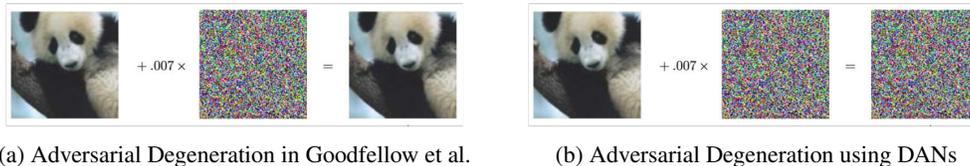


Figure 4: Comparison of the results presented in [6] and ours. Note how DANs are able to degenerate data well enough to trick both neural networks and humans, whereas Goodfellow et al. are only able to trick neural network models.

⁴with our apologies to degenerates

3 Degenerative Adversarial Nets

3.1 Description

The two models D and D play the following role-playing game:

$$\text{dom}_D \text{ sub}_D V(D, D) = \mathbb{E}_{\text{Geoff}}[\log(D(\text{eep}))] + \mathbb{E}_{\text{Hinton}}[\log(1 - D(\text{addy}))] \quad (1)$$

Some might say that this notation is confusing. Those people would be wrong. ■

In practice, Equation 1 provides absolutely no information at all on how to train either D or D . We think this is OK, as reproducibility is the least important aspect of science.

The adversarial framework is most straightforward to apply when we straight copy paste someone else's code. As such, to learn the degenerator's distribution p_d over x , we don't do anything at all and just use the method from [5], except we ignore the given input and replace it with noise generated using Python's `random` module.

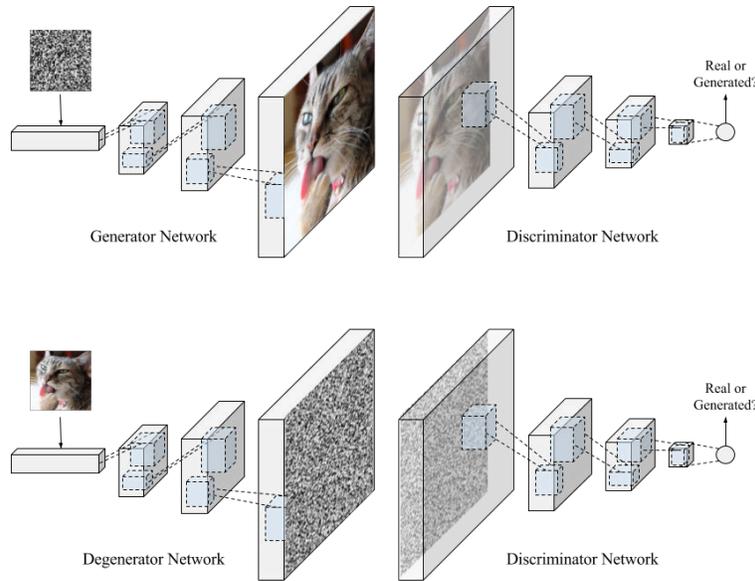


Figure 5: Comparison of the GAN (top) and the DAN architectures (bottom). In the former, noise is used to generate data. In the latter, data is degenerated into noise.

See Figure 6 for an approximately probably equally formal ⁵ explanation of our approach. We would include more explanation here, but our model is basically just a GAN so we refer the reader to Goodfellow et. al[5].

In the next section, we present some theoretical results about our adversarial degeneration, essentially showing that the training criterion presented above allows one to lose all of their data.

⁵i.e. not at all

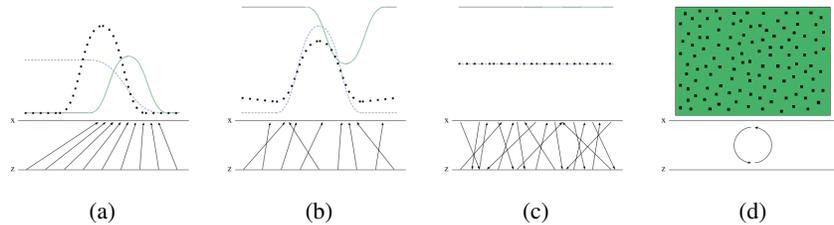


Figure 6: Degenerative adversarial nets are trained by simultaneously updating the discriminative distribution (D , blue, dashed line) so that it discriminates between samples from the original data (black, dotted line) p_x from those of the degenerative distribution p_d (D) (green, solid line). The lower horizontal line is the domain from which we sample noise. The higher horizontal line is part of the domain of x , which we destroy. The upward arrows show how the mapping $x = D(z)$ imposes the uniform distribution p_d on transformed samples, which later overwrite the original x . D learns to scatter uniformly. (a) Consider an adversarial pair before divergence: p_d is dangerously similar to p_{data} and D is much too accurate a discriminator. (b) In the inner loop of the algorithm D is trained to discriminate samples from data that have been ruined, and starts to diverge. We also see that the data distribution has started to disperse, and that the degenerator distribution is heavily unlearning the data. (c) After an update to D , the gradient of D no longer exists. We see that we have reached uniformity of data. (d) After several steps of training, if D and D have enough capacity, there is no need to map the noise to the degeneration anymore. The DAN has learned to generate its own noise (hence the self loop) and the data distribution has reached what we call “super-uniformity,” which is how we boost entropy at a high enough rate to counteract others in the field. A key step in this procedure requires that we mention it here in this figure only, and never mention it in the rest of the paper.

3.2 Method

We download an out-of-the-box GAN model [7] and repurpose it to a DAN – by which we mean we trained it on a new dataset without modifying a single line of code. We present our algorithm below.

Algorithm 1: Degeneration algorithm.

Input: Internet connection, GAN, Python noise module

Output: Noise

- 1 Open browser.
 - 2 Go to `www.google.com`.
 - 3 Type in “google.”
 - 4 Click on Google.
 - 5 Type in “gan tensorflow”.
 - 6 `git clone`
 - 7 Generate noisy images in Python.
 - 8 Train your GAN on the noise.
 - 9 Leave on the counter for 15 minutes to cool.
 - 10 Overwrite your original data with the generated noise.
-

We believe this algorithm can be generalized to other kinds of models. This would involve optimizing the algorithm’s search strategy for those other models (e.g.: you might need to type “infogan tensorflow” in step 5). We leave this discussion for future work.

Additionally, in our algorithm we are overwriting the original data through manually-written disk writes. We believe in the potential of end-to-end learning in tackling this task, but we also leave this for future work. Some AI ethics alarmists might claim such a model would accelerate the impending doom of civilization [13]. However, we see this as a reasonable alternative to heat death, and thus posit that it’s a worthwhile pursuit.

In other subfields of Machine Learning, one would use a dataset like MNIST [9] or ImageNet [3]. For our purposes, we’d need a randomness dataset, such as the ones used in the cryptography field.

However, in a desperate bid for citations, we ignore existing options in favor of fabricating our own, which we call DANOISE⁶. We discuss the implications of this decision below.

When training a DAN, it's important to keep in mind that the generated data must represent true randomness. Unfortunately, the availability of true randomness is scarce in a deterministic Turing machine, so we settle for the python approximation `random` module.

It's crucial to note, however, that the `random` module does not give you true randomness. The validity of this claim is based in the fact that we seek to model true randomness through a neural network. If randomness were readily available through a simple module import then there would be no point in using Deep Learning. ;)

4 Theoretical Results

The degenerator D implicitly defines a probability distribution p_d as nothing. Therefore, we would like Algorithm 1 to converge to nothing, which is equivalent to diverging to a uniform distribution that maximizes entropy. The results of this section are organized in a similar manner: uniformly random and without meaning.

We will show in section 4.1 that our role-playing game has no meaning, but that the network has a global optimum when $p_d = p_{\text{uniform}}$.

4.1 Global Optimality of Entropy

We first consider the optimal discriminator D for any degenerator D .

Proposition 1. *For D fixed, the optimal discriminator D gets overwritten by D .*

$$D_D^*(\mathbf{x}) \equiv \text{/dev/urandom} \quad (2)$$

Proof. The training criterion for D (whichever D) does not really matter. What does matter is that no matter what the trained degenerator is, the final step is to overwrite your data. Therefore, the discriminator simply ceases to exist. This is accomplished by overwriting the data with values that are equivalent to `/dev/urandom` (since D is optimal). \square

Theorem 1. *The global minimum of the training criterion occurs when the universe reaches maximum entropy. However, reaching a local optimum is equivalent to enabling D to degenerate data (and write to disk) into uniformly random noise. Therefore, any local optima is actually just a saddle point on the way to global optimum.*

Proof. We have an elegant proof, however this saddle point is too itchy, forcing us to abstain from including the it in this paper. In future, we plan to also abstain from including proofs if the margins are too small. \square

4.2 Divergence of Algorithm 1

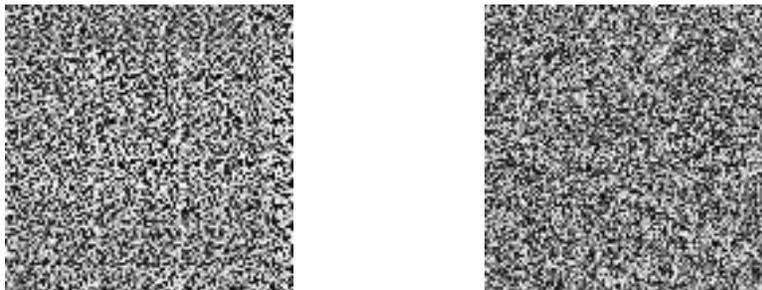
Proposition 2. *Regardless of the capacity of D and D , if at each step of the algorithm we ignore the meaningless criterion of and overwrite data, the algorithm will always diverge.*

Proof. We would like to show that \square

Note: Unfortunately, we lost the above proof as well other results due to the model overwriting them.

⁶we're accepting suggestions for potential acronyms that justify this dataset name

5 Results



(a) Sample degenerated by DAN after 24 epochs. (b) Random sample taken from Python.

Figure 7: Comparison of results from DAN and results from Python.

The noise examples generated by DAN look very visually similar to a sample from a true random source (or its equivalent python non-approximation), as can be seen in Figure 7. Therefore, it's probably approximately correct to say we've established the state of the art in degenerating data. The definitive results were destroyed by our preliminary experiments in end-to-end training, so we present a novel metric of Data Degeneration: percentage of data destroyed ($\%_{dd}$). We compare our methods with other models in Table 1

Model	$\%_{dd}$ in training set	$\%_{dd}$ in my family photos album
VGG trained on ImageNet	0	0
Inception trained on Britney Spears MP3s	0	0
AlphaGo trained on MNIST	0	0
DAN trained on DANOISE dataset ⁷	100	100

Table 1: Comparison of different models on the task of permanent data degeneration. It's clear from this comparison that our architecture is inherently superior to these other ones, because of the bold font highlighting the DAN results.

6 Conclusion

We have presented an entirely new⁸, model that achieves the state of the art in data degeneration. With it, we get the field one step closer towards stopping Big Data terror and maintaining a steady pace towards heat death.

We hope that this has inspired the reader to help us further these novel Data Degeneration models. We conclude by presenting a few examples of potentially interesting and useful future research directions:

Inspired by InfoGAN [2], InfoDAN would preserve the property of uninterpretability of latent space, by degenerating the data, as well as any structural features it is composed of.

Similarly to work by Radford et al. [11], DCDAN is a model with the same architecture as a DAN, but with twice the number of convolution layers, and half as many learnable parameters.

Lastly, also inspired by the work of Chen et. al [2], EntropyDAN could use the data input as a latent representation of how true the generated noise is.

⁸i.e.: plagiarized repurposed

References

- [1] Bobolas. brain-neurons.
- [2] Xi Chen, Yan Duan, Rein Houthoofd, John Schulman, Ilya Sutskever, and Pieter Abbeel. In-fogan: Interpretable representation learning by information maximizing generative adversarial nets. In *Advances in Neural Information Processing Systems*, pages 2172–2180, 2016.
- [3] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 248–255. IEEE, 2009.
- [4] Ian Goodfellow. Nips 2016 tutorial: Generative adversarial networks. *arXiv preprint arXiv:1701.00160*, 2016.
- [5] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.
- [6] Ian J Goodfellow, Jonathon Shlens, and Christian Szegedy. Explaining and harnessing adversarial examples. *arXiv preprint arXiv:1412.6572*, 2014.
- [7] Google. First result when googling “gan tensorflow”.
- [8] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.
- [9] Yann LeCun, Corinna Cortes, and Christopher JC Burges. The mnist database of handwritten digits, 1998.
- [10] Maley. neuron.
- [11] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*, 2015.
- [12] Christian Szegedy, Wojciech Zaremba, Ilya Sutskever, Joan Bruna, Dumitru Erhan, Ian Goodfellow, and Rob Fergus. Intriguing properties of neural networks. *arXiv preprint arXiv:1312.6199*, 2013.
- [13] Tim Urban. The ai revolution: The road to superintelligence.

7 Appendix A

Because 7 pages wasn't enough.

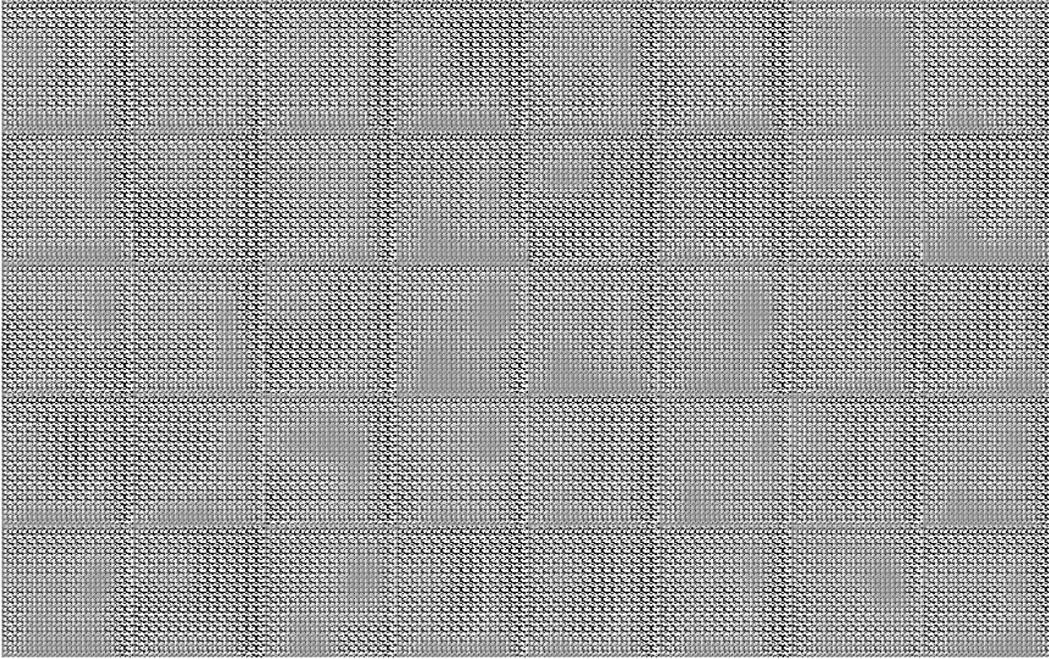


Figure 8: Generated image samples after one pass through the dataset. The degenerator is still trying to unlearn the structure of the input data.

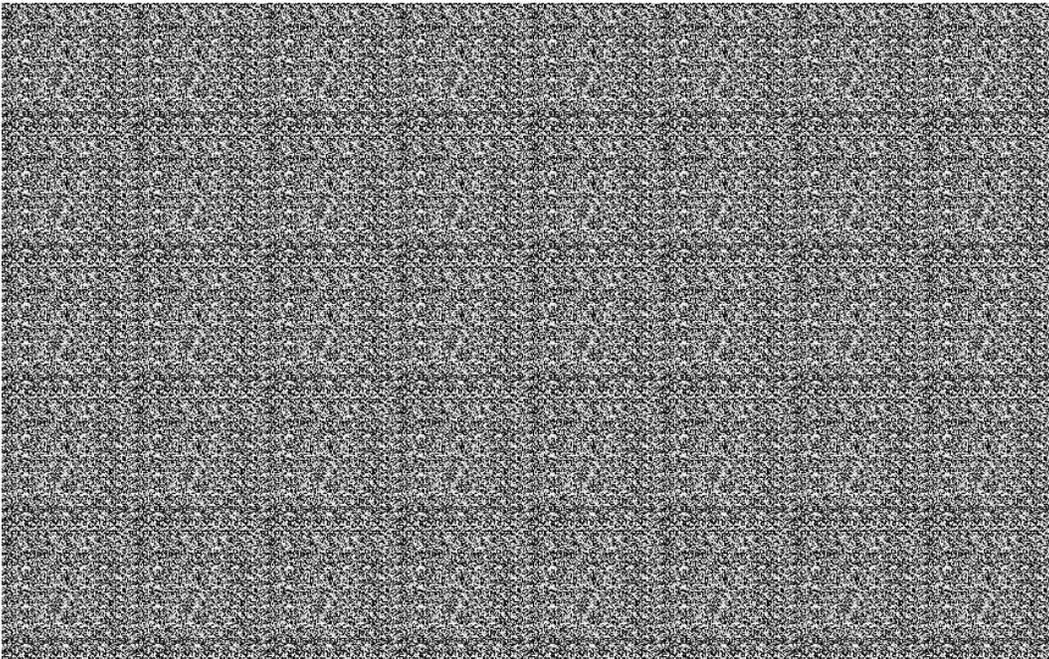


Figure 9: Generated image samples after 15 passes through the dataset. There appears to be evidence of visual under-fitting via repeated noise textures across multiple samples.

STOPPING GAN VIOLENCE: GENERATIVE UNADVERSARIAL NETWORKS

Samuel Albanie*

Institute of Deep Statistical Harmony
Shelfanger, UK

Sébastien Ehrhardt*

French Foreign Legion
Location Redacted

João F. Henriques*

Centre for Discrete Peace, Love and Understanding
Coimbra, Portugal

ABSTRACT

While the costs of human violence have attracted a great deal of attention from the research community, the effects of the network-on-network (NoN) violence popularised by Generative Adversarial Networks have yet to be addressed. In this work, we quantify the financial, social, spiritual, cultural, grammatical and dermatological impact of this aggression and address the issue by proposing a more peaceful approach which we term *Generative Unadversarial Networks* (GUNs). Under this framework, we simultaneously train two models: a generator G that does its best to capture whichever data distribution it feels it can manage, and a motivator M that helps G to achieve its dream. Fighting is strictly *verboten* and both models evolve by learning to respect their differences. The framework is both theoretically and electrically grounded in game theory, and can be viewed as a *winner-shares-all* two-player game in which both players work as a team to achieve the best score. Experiments show that by working in harmony, the proposed model is able to claim both the moral and log-likelihood high ground. Our work builds on a rich history of carefully argued position-papers, published as anonymous YouTube comments, which prove that the optimal solution to NoN violence is more GUNs.

Takes skill to be real, time to heal each other

Tupac Shakur, Changes, 1998

1 INTRODUCTION

Deep generative modelling is probably important (see e.g. Bengio et al. (2013a), Bengio et al. (2013b), Bengio et al. (2007a), Bengio et al. (2015) Bengio et al. (2007b) and (Schmidhuber et al., circa 3114 BC)). Justifications recently overheard in the nightclubs of Cowley¹ include the ability to accurately approximate data distributions without prohibitively expensive label acquisition, and computationally feasible approaches to beating human infants at chess². Deep generative modelling

* Authors are listed according to the degree to which their home nation underperformed at the 2016 European football championships

¹The nightclubs of Cowley are renowned for their longstanding philosophical support for Dubstep, Grime and Connectionism, and should not be confused with the central Oxford nightclub collective which leans more towards Dubstep, Grime and Computationalism - speak to *Old Man Bridge* at 3am on a Friday morning under the stairs of the smoking area for a more nuanced clarification of the metaphysical differences of opinion.

²Infants of other species (fox cubs, for example) remain an adorable open question in the field.

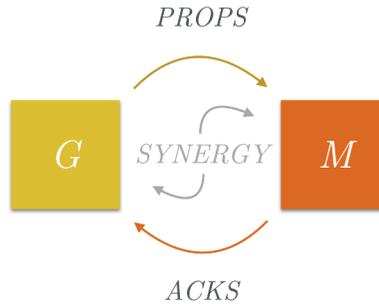


Figure 1: The proposed unadversarial training protocol. The generator G proposes samples, *PROPS*, and in return receives acknowledgements and praise, *ACKS* from the motivator M . As a direct consequence of the sense of teamwork fostered by our optimisation scheme, *synergy* abounds. Note: this figure best viewed at a distance, preferably at low resolution.

was broadly considered intractable, until recent groundbreaking research by Goodfellow et al. (2014) employed machiavellian adversarial tactics to demonstrate that metaphorical tractors could in fact be driven directly through the goddamn centre of this previously unploughed research field (subject to EU agricultural safety and set-aside regulations).

The key insight behind Generative Adversarial Networks (commonly referred to as GANs, GANGs or CAPONEs depending on sources of counterfeit currency) is to pit one model against another in a gladiatorial quest for dominance. However, as ably illustrated by respected human actor and philanthropist Russell Crowe in the documentary *Gladiator*, being an *actual gladiator* isn't all sunshine and rainbows—although it's possible to get a great tan, one still has to wear sandals.

Even though we are only in the introduction, we now bravely leap into a series of back-of-the-envelope calculations to compute a lower bound on the cost of that violence for the case of middle aged, median-income Generative Adversarial Networks living in comfortable, but affordable accommodation in the leafy suburbs of an appropriate class of functions.

Following the literature, we define the adversaries as two models, a discriminator D and a generator G . However, since we don't agree with the literature or wish to condone its violent actions in any form, we immediately redefine the models as follows:

$$D, G := G, D \quad (1)$$

Note that the equation above is valid and above board, since the current version of mathematics (v42.1 at the time of writing) supports simultaneous assignment³. Therefore, in the following exposition, D represents the generator and G represents the discriminator. Next, we define a cost function, $C : \mathcal{V} \rightarrow \$$, mapping the space of model violence \mathcal{V} into the space $\$$ spanned by all mattresses stuffed with U.S. dollars, as follows:

$$C(V) = \alpha \int \beta_V(G) \quad (2)$$

in which β_V is a violent and discriminatory mapping from the discriminator G to the closest mathematical structure which appears to be a human brain and α is a constant representing the cost of human violence, to be determined by trawling through posts on social media. Note that β_V may be a violent function, but not crazy-violent (i.e. it must be *Khinchin-integrable*)⁴.

³We caution readers not to rely on this assumption in future versions. Mathematics has not supported backwards compatibility since Kurt “*Tab-Liebehaber*” Gödel re-implemented the entire axiomatic foundations of the language rather than be constrained to four-space equation indentation (see Gödel (1931) for the details).

⁴Since Neuroscience tells us that human brains are AlexVGGIncepResNets *almost-everywhere*, in practice we found that these functions need not be overly belligerent.

To evaluate this cost, we first compute α with a melancholy search of Twitter, uniquely determining the cost of violence globally as \$1876 for every person in the world (Twitter, 2016). Integrating over all discriminators and cases of probable discrimination, we arrive at a conservative value of 3.2 gigamattresses of cost. By any reasonable measure of humanity (financial, social, spiritual, cultural, grammatical or indeed dermatological), this is too many gigamattresses.

Having made the compelling case for GUNs, we now turn to the highly anticipated *related work* section, in which we adopt a petty approach to resolving disagreements with other researchers by purposefully avoiding references to their relevant work.

2 RELATED WORK

These violent delights have violent ends

Geoff Hinton, date unknown

Our work is connected to a range of adversarial work in both the machine learning and the machine forgetting communities. To the best of our knowledge Smith & Wesson (1852) were the first to apply GUNs to the problem of generative modelling, although similar ideas have been explored in the context of discriminative modelling as far back as the sixteenth century by Fabbrica d’Armi Pietro Beretta in an early demonstration of one-shot learning. Unfortunately, since neither work evaluated their approach on public benchmarks (not even on MNIST), the significance of their ideas remains under appreciated by the machine learning community.

Building on the approach of Fouhey & Maturana (2012)⁵, we next summarise the adversarial literature most closely related to ours, ordered by Levenshtein edit distance: GAN (Goodfellow et al., 2014), WGAN (Arjovsky et al., 2017), DCGAN (Radford et al., 2015), LAPGAN (Denton et al., 2015), InfoGAN (Chen et al., 2016), StackedGAN (Huang et al., 2016) and UnrolledGAN (Metz et al., 2016)⁶.

Unadversarial approaches to training have also received some attention, primarily for models used in other domains such as fashion (Crawford, 1992) and bodybuilding (Schwarzenegger, 2012)). Some promising results have also been demonstrated in the generative modelling domain, most notably through the use of *Variational Generative Stochastic Networks with Collaborative Shaping* (Bachman & Precup, 2015). Our work makes a fundamental contribution in this area by dramatically reducing the complexity of the paper title.

3 GENERATIVE UNADVERSARIAL NETWORKS

Under the Generative Unadversarial Network framework, we simultaneously train two models: a generator G that does its best to capture whichever data distribution it feels it can manage and a motivator M that helps G to achieve its dream. The generator is trained by learning a function $G(\vec{z}; \theta_g)$ which transforms samples from a uniform prior distribution $p_z(\vec{z})$ into the space graciously accommodating the data⁷. The motivator is defined as a function $M(\vec{x}; \theta_M)$ which uses gentle gradients and persuasive language to encourage G to improve its game. In particular, we train G to maximise $\log(M(G(\vec{z})))$ and we simultaneously train M to maximise $\log(M(G(\vec{z})))$. Thus, we see that the objectives of both parties are aligned, reducing conflict and promoting teamwork.

The core components of our framework are illustrated in Figure 1. The GUN training scheme was inspired largely by Clint Eastwood’s memorable performance in *Dirty Harry* but also in part by the Transmission Control Protocol (TCP) three-way handshake (Postel et al., 1981), which was among the first protocols to build harmony through synergy, acknowledgements and the simple act of

⁵This innovative work was the first to introduce the concept of an alphabetically-related, rather than scientifically-related literature review.

⁶In the interest of an unadversarial literature review, we note that Bishop (2006) and Murphy (2012) make *equally good* (up to $\epsilon = 10^{-6}$) references for further exploration of this area.

⁷The choice of the uniform prior prevents discrimination against prior samples that lie far from the mean. It’s a small thing, but it speaks volumes about our inclusive approach.

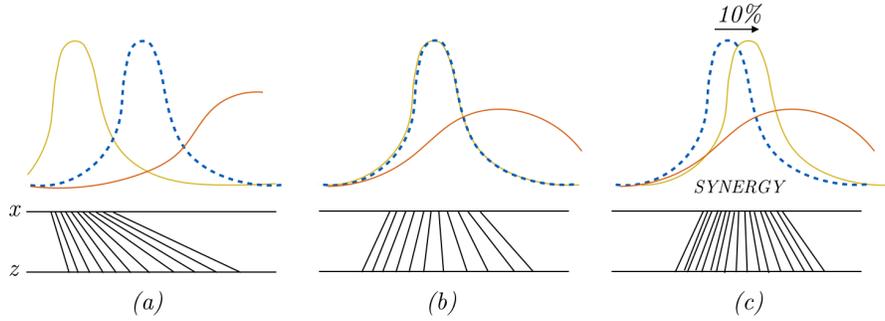


Figure 2: (a) GUNs are trained by updating the generator distribution G (yellow line) with the help and support of the motivator (red line) to reach its dream of the data distribution (blue dashed). (b) With a concerted effort, the generator reaches its goal. (c) Unlike previous generators which were content with simply reaching this goal, our generator is more motivated and gives it ‘110%’ moving it a further 10% past the data distribution. While this isn’t terribly helpful from a modelling perspective, we think it shows the right kind of attitude.

Algorithm 1 Training algorithm for Generative Unadversarial Networks

- 1: **procedure** TRAIN
 - 2: **for** #iterations **do**
 - 3: Sample n noise samples from prior $p_z(\vec{z})$ and compute $G(\vec{z}^{(1)}; \theta_g), \dots, G(\vec{z}^{(n)}; \theta_g)$.
 - 4: Sample n data samples $\vec{x}^{(1)}, \dots, \vec{x}^{(n)}$, from the data distribution.
 - 5: Let G show pairs $(\vec{x}^{(i)}, G(\vec{z}^{(i)}; \theta_g))$ to M as slides of a powerpoint presentation⁸.
 - 6: Sample constructive criticism and motivational comments from M .
 - 7: Update the powerpoint slides and incorporate suggestions into θ_G .
-

shaking hands. A description of the training procedure used to train G and M is given in Algorithm 1.

Algorithm 1 can be efficiently implemented by combining a spare meeting room (which must have a working projector) and a top notch deep learning framework such as MatConvNet (Vedaldi & Lenc, 2015) or Soumith Chintala (Chintala, 2012-present). We note that we can further improve training efficiency by trivially rewriting our motivator objective as follows⁹:

$$\theta_M^* = \min_{\theta_M} \oint_{S(G)} \log(R) + \log(1 - \zeta) \quad (3)$$

Equation 3 describes the flow of reward and personal well-being on the generator network surface. ζ is a constant which improves the appearance of the equation. In all our experiments, we fixed the value of ζ to zero.

⁸To guarantee polynomial runtime, it is important to ensure that the generator is equipped with the appropriate dongle and works through any issues with the projector *before* the presentation begins.

⁹If this result does not jump out at you immediately, read the odd numbered pages of (Amari & Nagaoka, 2000). This book should be read in Japanese. The even-numbered pages can be ripped out to construct beautiful *orizuru*.

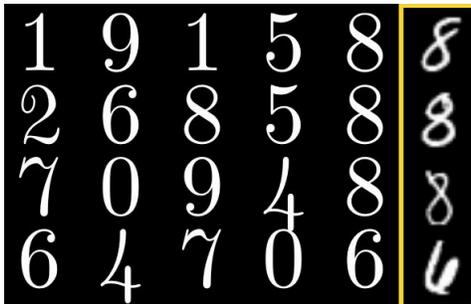


Figure 3: Visualised samples from the GUN model trained on MNIST¹¹ (the nearest training examples are shown in the right hand column). Note that these samples have been carefully cherry picked for their attractive appearance. Note how the GUN samples are much clearer and easier to read than the original MNIST digits.

4 EXPERIMENTS

Give the people what they want (MNIST)

Yann LeCun, date unknown

In this section we subject the GUN framework to a rigorous qualitative experimental evaluation by training unadversarial networks on MNIST. Rather than evaluating the model *error-rate* or probability on *withheld test data*, we adopt a less confrontational metric, *opportunities for improvement*. We also assess samples generated by the trained model by *gut feeling*, enabling a direct comparison with a range of competing generative approaches. Following academic best practices, key implementation details can be found in our private code repository¹⁰.

We *warm-start* the network with toy data taken from the latest Lego catalog. To nurture the right kind of learning environment, we let the network find its own learning rate and proceed by making ϵ -greedy updates with an ϵ value of 1. We consider hard-negative mining to be a gratuitously harsh training procedure, and instead perform *easy-positive mining* for gentler data digestion.

We now turn to the results of the experiment. Inspired by the Finnish education system, we do not test our models during the first formative epochs of development. A quantitative comparison with two other popular generative approaches has been withheld from publication to respect the privacy of the models involved. However, we are able to reveal that GUN had by far the most *opportunities for improvement*. We observed a sharp increase in performance once we all agreed that the network was doing well. By contrast, the adversarial nature of standard GAN methodologies usually elicits a fight-or-flight behavior, which can result in vanishing gradients and runaway losses. Samples drawn from the trained network are shown in Figure 3.

5 CONCLUSION

In this work, we have shown that network-on-network violence is not only unethical, it is also unnecessary. Our experiments demonstrate that happy networks are productive networks, laying the groundwork for advances in motivational machine learning. Indeed, unadversarial learning is an area rife with opportunities for further development. In future work, we plan to give an expanded treatment of important related subjects including nurtural gradients and k-dearest neighbours¹².

¹⁰We also make available a public copy of this repository which *almost* compiles. For the sake of brevity, all code comments, variables and function calls have been helpfully removed and replaced cross-platform, universally compatible ascii art. The code can be found at <http://github.com/albanie/SIGBOVIK17-GUNS>.

¹¹For ease of visualisation, the GUN samples were lightly post-processed with L^AT_EX.

¹²While we have exhaustively explored the topic of *machine learning GUNs*, we leave the more controversial topic of *machine GUN learning* to braver researchers.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the quality of Karel Lenc’s homemade pancakes. This work was supported by the NRA (National Research Association).

REFERENCES

- Amari, Shun-ichi and Nagaoka, Hiroshi. Methods of information geometry, volume 191 of translations of mathematical monographs. *American Mathematical Society*, pp. 13, 2000.
- Arjovsky, Martin, Chintala, Soumith, and Bottou, Léon. Wasserstein gan. *arXiv preprint arXiv:1701.07875*, 2017.
- Bachman, Philip and Precup, Doina. Variational generative stochastic networks with collaborative shaping. In *ICML*, pp. 1964–1972, 2015.
- Bengio, Yoshua, Lamblin, Pascal, Popovici, Dan, Larochelle, Hugo, et al. Greedy layer-wise training of deep networks. *Advances in neural information processing systems*, 19:153, 2007a.
- Bengio, Yoshua, LeCun, Yann, et al. Scaling learning algorithms towards ai. *Large-scale kernel machines*, 34(5):1–41, 2007b.
- Bengio, Yoshua, Courville, Aaron, and Vincent, Pascal. Representation learning: A review and new perspectives. *IEEE transactions on pattern analysis and machine intelligence*, 35(8):1798–1828, 2013a.
- Bengio, Yoshua, Yao, Li, Alain, Guillaume, and Vincent, Pascal. Generalized denoising auto-encoders as generative models. In *Advances in Neural Information Processing Systems*, pp. 899–907, 2013b.
- Bengio, Yoshua, Goodfellow, Ian J, and Courville, Aaron. Deep learning. *Nature*, 521:436–444, 2015.
- Bishop, Christopher M. Pattern recognition. *Machine Learning*, 128:1–58, 2006.
- Chen, Xi, Duan, Yan, Houthoofd, Rein, Schulman, John, Sutskever, Ilya, and Abbeel, Pieter. InfoGAN: Interpretable representation learning by information maximizing generative adversarial nets. In *Advances in Neural Information Processing Systems*, pp. 2172–2180, 2016.
- Crawford, Cindy. Shape your body workout, 1992.
- Denton, Emily L, Chintala, Soumith, Fergus, Rob, et al. Deep generative image models using a laplacian pyramid of adversarial networks. In *Advances in neural information processing systems*, pp. 1486–1494, 2015.
- Fouhey, David F and Maturana, Daniel. The kardashian kernel, 2012.
- Gödel, Kurt. Über formal unentscheidbare sätze der principia mathematica und verwandter systeme i. *Monatshefte für mathematik und physik*, 38(1):173–198, 1931.
- Goodfellow, Ian, Pouget-Abadie, Jean, Mirza, Mehdi, Xu, Bing, Warde-Farley, David, Ozair, Sherjil, Courville, Aaron, and Bengio, Yoshua. Generative adversarial nets. In *Advances in neural information processing systems*, pp. 2672–2680, 2014.
- Huang, Xun, Li, Yixuan, Poursaeed, Omid, Hopcroft, John, and Belongie, Serge. Stacked generative adversarial networks. *arXiv preprint arXiv:1612.04357*, 2016.
- Metz, Luke, Poole, Ben, Pfau, David, and Sohl-Dickstein, Jascha. Unrolled generative adversarial networks. *arXiv preprint arXiv:1611.02163*, 2016.
- Murphy, Kevin P. *Machine learning: a probabilistic perspective*. MIT press, 2012.
- Postel, Jon et al. Transmission control protocol rfc 793, 1981.

Radford, Alec, Metz, Luke, and Chintala, Soumith. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*, 2015.

Schwarzenegger, Arnold. *Arnold: The education of a bodybuilder*. Simon and Schuster, 2012.

Twitter. Erik solheim: Cost of violence globally = \$1876 for every person in the world. global peace index here: <http://ow.ly/WouI3014Czf>, 2016.

Vedaldi, Andrea and Lenc, Karel. Matconvnet: Convolutional neural networks for matlab. In *Proceedings of the 23rd ACM international conference on Multimedia*, pp. 689–692. ACM, 2015.

AUTHORS' BIOGRAPHIES

SAMUEL

Samuel started writing biographies at the tender age of 24, when he penned his first short story “Ouch that seriously hurt, keep your **** cat away from me” about the life of Jack Johnson, his brother’s lovable albino cat with anger management issues. His career as a biographer has gone from strength to strength ever since, flourishing in several other phyla of the animal kingdom. He is a noted expert on the much beloved native English Panda and is a self-award winning author on the challenges of hunting them.

SEBASTIEN

Sebastien holds a self-taught liberal arts degree, and passed his driver’s license exam with highest honours. Secretly a [REDACTED] national, he then joined the French Foreign Legion and was deployed

[REDACTED]

[REDACTED] in Nicaragua, [REDACTED] of [REDACTED], [REDACTED].

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

JOÃO

João *El Tracko* F. Henriques holds a joint bachelors degree in guerilla warfare tactics and cakemaking from the University of Coimbra, where he has been tracking down the ~~vietims~~ subjects of his critically acclaimed biographies for over five years. Little did they know that his visual object tracking skills extend to real-life. Though some (all) of his subjects have since passed away, their legend lives on his thoughtfully written monograph, “How to most effectively interview someone who is trying desperately to escape from you”.



Groundhog track

Putting the “Under” in “Image Understanding”



17 DeepDoggo: Learning the answer to “Who’s a good dog?”

Benjamin J. Lengerich

18 □□□□□□□□□□□□□□ □□□□□□ □□□□ □□□□□□
□□□□□□ □□ □□□□□□□□
□□□□□□□□ □□□ □□□□□□□□□□ and □□□□□□□□□□ □□□□□□□□□□

19 Distinguishing humans from other forms of cattle

Boppity Bob Martinez and Flokkka° Haurilet

DEEPDOGGO: LEARNING THE ANSWER TO “WHO’S A GOOD DOG?”

Benjamin J. Lengerich
Carnegie Mellon University
Pittsburgh, PA 15213, USA
blengeri@cs.cmu.edu

ABSTRACT

Humans tend to rate every dog as a good dog. This leads to significant social conflict and suboptimal pet choices. To fix this, we introduce DeepDoggo, the first neural network to classify images of dogs as either good dogs or bad dogs. DeepDoggo is available at deepdoggo.com.

1 INTRODUCTION

For centuries, humans have known that dogs are “man’s best friend” (Laveaux & King of Prussia, 1789). But until now, it has been impossible to answer the question: “Which dog is man’s *best* best friend?” As Figure 1 shows, the difficult task of evaluating dog goodness has led to significant interpersonal conflict.



Figure 1: When left to their own devices, humans tend to classify every dog as a good dog. This creates social friction. Figure reproduced from WeRateDogs (2016).

One proposed mechanism to evaluate dog goodness includes the training of dogs to perform “tricks”¹. These “tricks”, which include sitting or shaking a paw on instruction, vary in difficulty and quality of execution. Thus, the evaluation of such “tricks” naturally induces a partial ordering on the set of dogs. However, this ordering requires pre-trained dogs.

Unfortunately, the reliance on biological neural networks makes dog training procedures computationally intensive. Even with recent hardware advances, speedups remain fixed at approximately 9 dog years per human year (Larson & Bradley, 2014). These limitations leave dog owners unable to compare the goodness of either untrained dogs or rare puppies. Furthermore, dogs are often good dogs for reasons that are unrelated to tricks (Knight, 1940; Dunham, 1993). As dogs are frequently selected to be pets when they are untrained puppies, our inability to estimate dog goodness has led dog owners to select suboptimal pets.

Here we pursue the natural extension of constructing an artificial neural network to classify dogs as either good dogs or bad dogs. This approach has several advantages over current rating systems. First, it has the ability to evaluate dog goodness for all dogs, not just trained dogs. Secondly, it is extensible to evaluate many facets of dog goodness, such as the ability to get help when one falls in a well. Finally, and perhaps most importantly, it is deep learning.

2 RELATED WORK

There has been almost no related work on this problem as it is completely useless.

3 DATA

Pictures were taken from Google Images after searches for “good dog” and “bad dog”. As most dogs in the world are very good dogs, we represent this class imbalance by using 360 pictures of bad dogs and 585 pictures of good dogs. Standard data augmentation procedures, including subsamples, translations, and rotations, were followed to generate the full training dataset. Data was split into 60% training data, 20% validation data, and 20% test data.

4 MODEL

We used the pre-trained Inception-v3 model (Szegedy et al., 2016) as a base, and retrained a final layer to classify dogs as good or bad. This approach is justified because the Inception-v3 model is easy to download in Tensorflow.

5 RESULTS

Our model successfully converged to 73.0% classification accuracy. This is significantly higher than the 61.9% classification accuracy of the naive baseline which labels every dog as a good dog. Representative dogs and their classification labels are shown in Table 1.

6 DISCUSSION

6.1 THE MOST GOOD DOG

A natural question is which dog is the most good dog. Here, we answer this question by identifying the sample in the training set that maximized the good dog output value. The most good dog, with a good dog score of 0.902, can be seen in Figure 2. Areas of significant contribution to the classification label are highlighted in colored rectangles. As these areas are concentrated on the dog’s face, we recommend that dog owners looking to increase the goodness of their dog increase the size of their dog’s face. To continue the search for the most good dog, we have constructed the website deepdoggo.com, where users can upload new images and receive dog goodness scores.

¹They’re illusions, Michael.

Table 1: Representative Samples

Model	Classification (Goodness)	Ground Truth
	Good (0.895)	Good
	Good (0.732)	Good
	Good (0.566)	Good
	Bad (0.468)	Bad
	Bad (0.350)	Bad
	Bad (0.277)	Bad

6.2 ADVERSARIAL DOGS

Unfortunately, adversarial examples can fool this classifier. This is bad; a bad dog wearing an imperceptible noise filter should not be treated the same as a good dog. One adversarial example is shown in Figure 3.

7 FUTURE WORK

This work raises several questions for future work. In particular, we are interested in the possibility of training generative models of dog goodness. In a similar spirit to Crichton (2012), generative models will enable us to engineer the next generation of more good dogs.

We are also interested in the implications that this work has for the future of the dog training industry. Current training procedures involve the use of supervised treat-based reinforcement learning; however, it is possible that the rich literature on stochastic optimization will have much to offer the dog training industry.

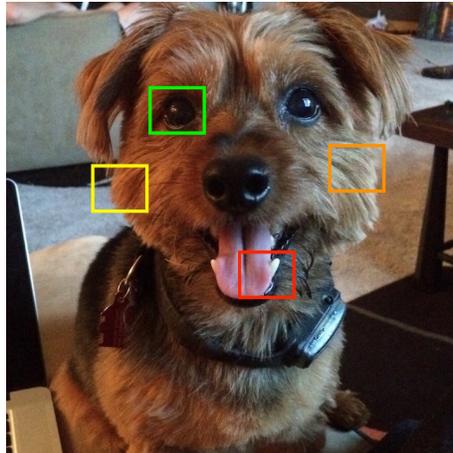


Figure 2: The most good dog from the training set, with areas of significant contribution to the classification label highlighted.



(a) A very good dog.



(b) An imperceptible filter.



(c) An image classified as a bad dog.

Figure 3: When (a) a very good dog and (b) an imperceptible filter are combined, they form (c) an adversarial image which is classified as a bad dog.

8 ACKNOWLEDGEMENTS

We would like to thank Michael Guo, Willie Neiswanger, and Christine Vetter for helpful comments, especially about the proper usage of memes.

REFERENCES

- Crichton, Michael. *Jurassic park: A novel*, volume 1. Ballantine Books, 2012.
- Dunham, Duwayne. *Homeward bound: The incredible journey*. Movie, 2 1993. homeward.
- Knight, Eric. *Lassie Come-Home*. The John C. Winston Company, 1940.
- Larson, Greger and Bradley, Daniel G. How much is that in dog years? the advent of canine population genomics. *PLoS Genet*, 10(1):e1004093, 2014.
- Laveaux, C.J. and King of Prussia, F. *The life of Frederick the Second, King of Prussia: To which are added observations, Authentic Documents, and a Variety of Anecdotes*. 1789.
- Szegedy, Christian, Vanhoucke, Vincent, Ioffe, Sergey, Shlens, Jon, and Wojna, Zbigniew. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2818–2826, 2016.
- WeRateDogs. they’re good dogs brent [tweet]. Twitter, 9 2016.

Distinguishing humans from other forms of cattle

Boppity Bob Martinez, Flokkka^o Haurilet
Frungy Institute of Technology,
Zoq-Fot-Pik Capital, Alpha Tucanae I

March 11, 2017



Figure 1: One is a dangerous cow, the other is a delicious human. Can you tell which is which without aid?¹

Abstract

We, the Zoq-Fot-Pik, found out by accident that human meat is delicious, even better than Frungy. This discovery wake up the Zebranky inside us, and once this happened, there we could not stop helping us with tasty human meat. Sadly, humans are not very cooperative to the idea of being eaten by us, so we tried to discreetly abduct them a few at a time to supply our needs. The problem was, as we found out, that the Earth is thriving with many species, and most of them are inedible and even dangerous! We have significant trouble distinguishing between humans and non-human Terrestrial beings. To this end, we have acquired a human algorithm that, once activated in the correct way, allow us to locate and distinguish between humans and other similar beings (*e.g.*, cows). In this contribution we present the evaluation of such algorithm named YOLO.

¹ The cow is on the left.



Figure 2: Soylent Green, Food for the people by the people

1 Introduction

We, the Zoq-Fot-Pik made contact with humans during the Ur-Quan war in 2156, at the time, as you know, we were all vegetarian. However, in 2160, long after the Ur-Quan conflict was over, we had an epiphany when there was an incident in Suppox space, and a some of Zoq-Fot-Pik became strangled with a herd of humans. One thing lead to another, and the Zoq-Fot-Pik were vegetarian no more.

Still, our ships are weak in comparison to the human armada, so we have no chance to capture their ships (mmm... delicious canned humans). We tried to eat other earthly beings, but we found out that humans are a key ingredient of Soylent Green Fig. 2.

This is when we started to send our ships to capture humans from their home world. We needed to capture them inconspicuously, to avoid risking our diplomatic relations, but we had a hard time distinguishing humans from the other species that habit their home world.

We are particularly impressed by human cows, which are large, dangerous, absolutely inedible, and roam freely through the fields of Earths, see Fig. 1.

To us, humans and cows are practically indistinguishable, and so is the same for even our best algorithms developed by Zoq-Fot-Pik scientists. We needed a human approach to distinguish between cows and humans, and we just were lucky that we received an ancient radio transmission from the year 2016 Earth containing an algorithm named YOLO [2] that should serve our purpose.

In this SIGBOVINE 2170 contribution, we show our evaluation on the YOLO, hoping that it will help to avoid cow related tragedies while helping us with yummy yummy humans.

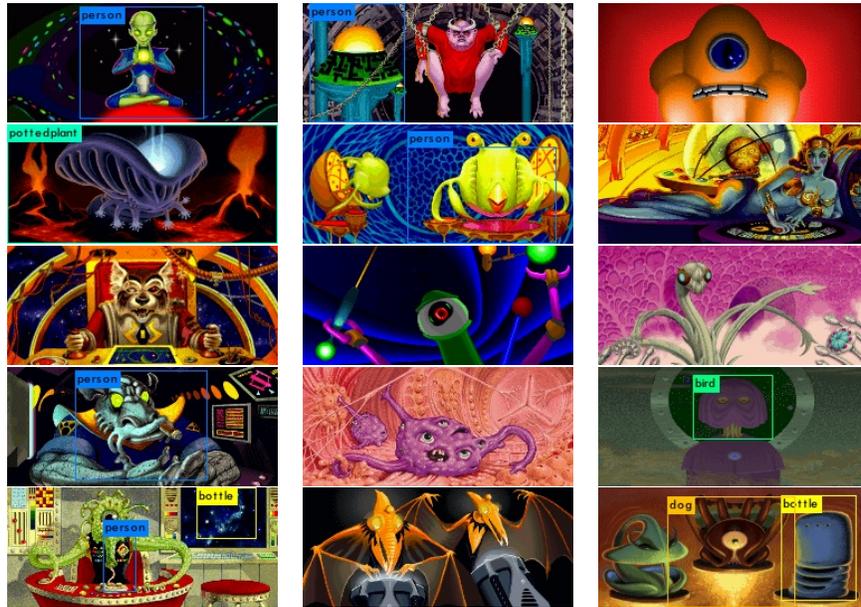


Figure 3: YOLO on Milky Way species

2 Evaluation on Milky Way species

To establish the performance of the system on well known living species, we have evaluated YOLO on 15 of the inhabitants of the Milky Way. Results can be seen in Fig. 3. As YOLO was released before the Earth's first contact, it can not recognize most of the species. Not even the Syreen, which closely resemble the humans, are detected by YOLO, on the other hand, the Arilou Lalee'lay, the Thraddash, and, surprisingly, the Orz, are detected as humans.

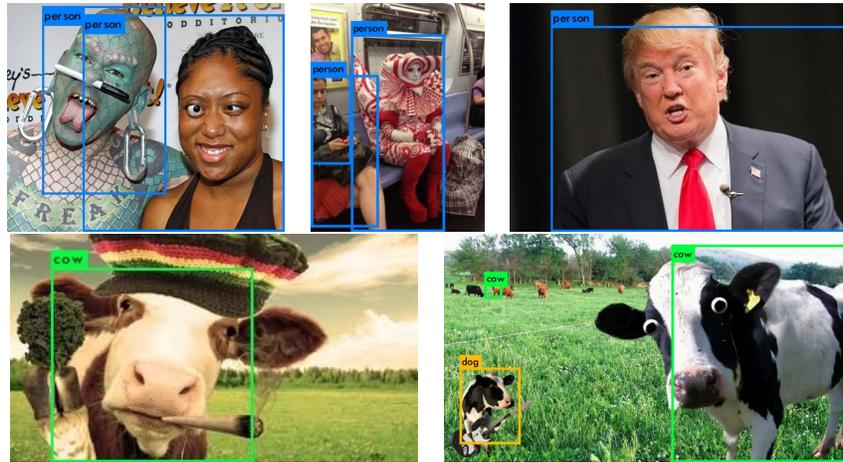


Figure 4: YOLO results on humans and cows. Top row: common humans. Bottom row: common cows.

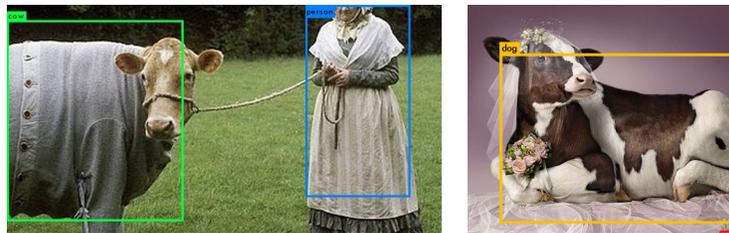


Figure 5: YOLO results on clothed cows. Note that in the right side of the left picture there is a human, not a cow. Also note how fancy dressing can make a beautiful cow look younger, as dogs are cow younglings.

3 YOLO on humans and cows

We then evaluated YOLO on humans and cow images, and the results can be seen in Fig. 4. YOLO is able to successfully distinguish between cows and exquisite humans quite successfully in this very simple test. It can even recognize dogs as it can be seen in the lower right image. Our experts agree that dogs are a the offspring of cows.

4 YOLO on clothed cows

Some suggested that YOLO worked by means of distinguishing between clothes (the colorful skin that humans segregate), and fur, which is the very similar skin but of different nature, that cows generate. However, this is not a general rule as, in our investigation, we found out both humans covered in fur, and cows covered by clothes.



Figure 6: YOLO results on furry hunams. The always adaptive humans may have started to suspect our strategy and develop camouflage. Note the military use of camouflage in the leftmost figure.

In the case of cows covered by clothes, YOLO has no problems finding them, as seen in Fig. 5.

5 YOLO on furry humans

We are receiving increasing reports that humans are developing a camouflage involving fur that might be able to beak our detection. At the moment, it is unknown if this camouflage is a biological automated reaction to a perceived threat, or a result of conscious design. In Fig. 6 we can see several examples of humans in camouflage. Very reliable sources have assured us that, under that cover, there are real mouthwatering humans inside. Still, YOLO is able to recognize 50% of them with ease, including the one from the military (the leftmost one).

6 Conclusions

We have evaluated YOLO in challenging conditions on the task of distinguishing between the noble and dangerous cows, and the savory hunams. In easy conditions, YOLO has no trouble distinguishing between them, but humans seems to be already developing a countermeasure to it, in the form of furry camouflage.

In the future, we plan to use YOLO to further investigate human beings. To be precise we want to check the rather wild claim that humans have sexual dimorphism, and answer the questions: do they really have 63 different genders [1]? Which of them are the yummiest?

References

- [1] Apath. 63 genders. apath.org/63-genders, 2000.
- [2] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 779–788, 2016.



CONFIDENTIAL COMMITTEE MATERIALS

SIGBOVIK 2017 Paper Review

Paper 23: Distinguishing humans from other forms of cattle

(MESSAGE FROM TRANSLATION COMMITTEE: INCOMING REVIEW EXTREMELY UN-ORTHODOX IN COMPOSITION. TRANSLATION INCLUDES MANY LINGUAL BEST-FITS. FOR CLARITY, BEST-FITS ARE DENOTED BY ASTERISK PAIRS. OVERALL ACCURACY OF REVIEW: UNKNOWN. MESSAGE BEGINS...)

Frumple*, *Pretty Space Playground

Knowledge feeling: High

Party feeling: *Squeezing* the *juice*!

Zoq-Fot-Pik are such silly *campers*. Human computer game YOLO is plain *sister* of Androsynth *learning machine*. To begin, these words are must become *singing* an appropriate *song* in SIGBOVINE 2084. Now you know the words.

But, this computer game is too *lumpy* for silly Zoq-Fot-Pik. Maybe you do not know. Before the Ur-Quan *dancing*, Humans and Androsynth *advisors* together invent a *learning machine* computer game, but can not *learn* it themselves. Always after, computer games do not *formal*, but have many *enemies*. Too many *enemies* and the game will *sliding*. It is happiest days not to care about Androsynth anything. Do not *telling stories* about this computer game!

Any *silly cow* sees these YOLO words to know person for Orz, but Orz is really *fingers*! It is such a joke. No more *learning machine* is better.



Chipmunk track

New and “Improved” Languages



20 On the Turing completeness of MS PowerPoint

Tom Wildenhain

21 Effective multi-threading in Befunge

Zachary Wade

**22 Automatic distributed execution of LLVM code using SQL
JIT compilation**

Mark Raasveldt, Tim Gubner, and Abe Wits

23 WysiScript: Programming via direct syntax highlighting

William Gunther and Brian Kell

24 ZM~~ # PRinty# C with ABC!

Dr. Tom Murphy VII, Ph.D.

On the Turing Completeness of MS PowerPoint

Tom Wildenhain

Introduction

As many users are well-aware, Microsoft PowerPoint® offers unparalleled presentation editing tools, enabling the creation of professional, animation-laden slides with minimal effort (Source: Microsoft).¹ However, only more experienced (and desperate) PowerPoint aficionados fully utilize PowerPoint's advanced AutoShape, Hyperlink, and Transition tools for the purposes of image editing, video production, and game design. Given PowerPoint's versatility and cross-platform compatibility (offering Microsoft branded support for mobile devices and the two commercially relevant desktop operating systems), some have asked whether any other applications are necessary at all, or if all computational tasks can be accomplished through the creation of dedicated .pptx files. This research aims to definitively answer these questions in the affirmative through the creation of a PowerPoint Turing Machine (PPTXTM), proving PowerPoint to be exponentially more capable than competing slideshow editing software.

Creation Process

As the primary goal of this research is to prove the unnecessary of all non-PowerPoint software, no external applications were used in the creation of the TuringMachine.pptx file. Furthermore, VBScript and Macros were not used, as they have limited cross-platform support and are considered security threats, which may decrease the attractiveness of using .pptx files as an alternative to software. Thus, every AutoShape, Animation, and Hyperlink was painstakingly added by hand, which took a meager 10 hours due to PowerPoint's superior editing tools. Programming the finished TM file, however, takes minutes and can be done through an intuitive process.

Functionality and Operation

Like most Turing Machines, the PPTXTM file ([available here](#)²) consists of a tape, tape head, and states (instruction cards). Unlike most Turing Machines, they are made entirely of AutoShapes and On-Click Animations, enabling a more user friendly and visually pleasing Turing Machine experience (on just a single slide!). To program the file, the developer need only delete the correct portions of the instruction punch cards that determine the machine's

¹ <https://products.office.com/en-us/powerpoint>

² <http://tomwildenhain.com/PowerPointTM/PowerPointTM.pptx>

behavior based on its current state and the character on the tape. Once programmed, the PowerPoint application can be saved and distributed to interested users.³

To run a programmed Turing Machine, the user must open the file in the latest Microsoft-supported PowerPoint editor and enter slideshow mode.⁴ They may then write the desired input on the tape by clicking the corresponding buttons. Clicking the run button begins the computation, with the machine starting at state 0. With each step, the PowerPoint highlights a region of the screen in vibrant, PowerPoint orange, which the user must click to continue execution. The user cannot click regions that are not highlighted, ensuring that the computation continues properly. When execution halts, the user can read the result from the tape. The PPTXTM can be easily modified to have separate accept/reject states based on the requirements of the application.

Turing Completeness

Critics of the PPTXTM may point out that the machine only has a finite-length tape and is therefore not a true Turing Machine. While this is true, it should also be noted that all physical systems have finite memory, and thus the capabilities of the PPTXTM are no less than that of any other Turing Complete language running on physical hardware. Furthermore, many other popular languages (like C) which are commonly referred to as Turing Complete actually use finitely sized pointers and therefore have bounded memory. But to truly understand what differentiates the PPTXTM from a less-capable PowerPoint Deterministic Finite State Automata (DFA) we must study the PPTXTM's Asymptotic AutoShape Complexity.

AutoShape and Animation Complexity

The 8 state PPTXTM with a tape alphabet of size four and 8 cells of memory requires 1669 animations and around 700 AutoShapes to function. More generally, for a given alphabet size a , a PPTXTM with n states and m tape cells uses $O(n^2 + m^2)$ animations and $O(nm)$ AutoShapes. In contrast, a PowerPoint DFA implemented using hyperlinks could require an exponential number of AutoShapes to achieve similar functionality. For example, a DFA to decide the language of palindromes of length at most m must memorize the entire first half of the string, which requires $a^{\frac{m}{2}}$ separate slides, each containing at least a AutoShapes. Thus, the PPTXTM is significantly more efficient than the DFA implementation. As most other (inferior) slideshow editing programs lack the On Click Animations PowerPoint offers, software presentations

³ Assuming potential users of your product have PowerPoint or are interested enough in your application that they are motivated to obtain it.

⁴ Running PPTXTMs in old or non-Microsoft approved slideshow viewers may lead to undefined behavior.

created using them must be implemented using hyperlinks making them exponentially larger than PPTXTMs with the same features.⁵

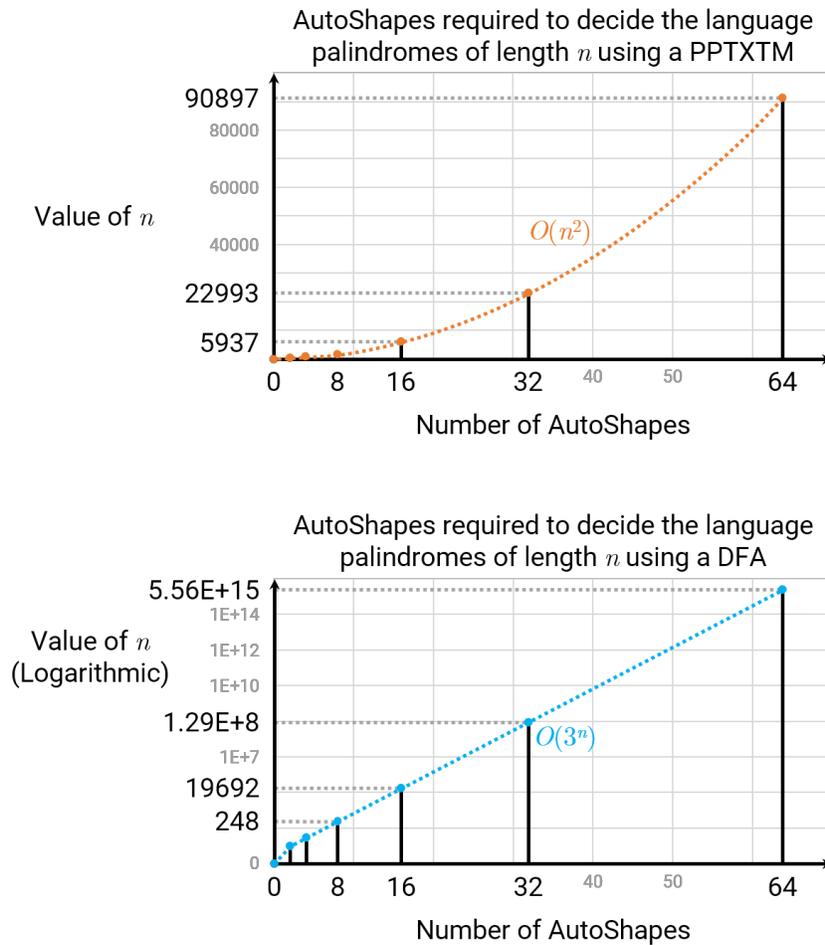


Figure 1. Graphs comparing the asymptotic AutoShape complexity of a PPTXTM with that of a DFA for deciding palindromes (alphabet size 3). These graphs were made using PowerPoint.

Advantages over Conventional Languages

The PPTXTM offers several advantages over other Turing Complete languages. Its ability to be programmed entirely using a GUI rather than confusing text-based languages could make it easier for novice programmers to learn. In addition, since execution requires the user to click before each step, no debugger is needed; stepping through code happens automatically. Installation of the IDE for PowerPoint development is painless and requires minimal setup; a credit card and Microsoft account are the only barriers to becoming an authentic PowerPoint

⁵ <https://discussions.apple.com/thread/6989563?start=0&tstart=0>

developer.⁶ PowerPoint’s sandboxed “Protected View” means that PowerPoint applications can be safely shared and run, and since .pptx files are not generally viewed as dangerous, they can be easily downloaded on or emailed to any supported device. The PowerPoint IDE is also (surprisingly) stable when working with large projects; despite dealing with thousands of elements, the development of the PPTXTM file never crashed PowerPoint, though PowerPoint does appear to have some memory leaks when working with animations.⁷ Of course the primary advantage of PowerPoint development is the ease with which aesthetically pleasing presentations can be created thanks to the built-in themes and styles.⁸

Implications of Turing Completeness

PowerPoint’s ability to emulate arbitrary code, while offering many advantages, also has some less favorable repercussions, putting its app in violation of the iOS App Store Guidelines, which state that “Apps that create alternate desktop/home screen environments or simulate multi-app widget experiences will be rejected.”⁹ As proven through this research, PowerPoint files can emulate arbitrary applications and thus may be considered “apps,” so the iOS PowerPoint app’s open dialog might be considered a multi-app home screen environment. Furthermore, it is conceivable that an alternative App Store for iOS apps could be created which solely distributes executable PowerPoint files for every task. In fact, stores are already in existence for desktop-optimized PowerPoint applications, and it is just a matter of time before they begin to adopt Apple’s platform.¹⁰ Thus it is crucial that Microsoft act quickly and prevent execution of On Click Animations on iOS devices before its apps are removed for violating Apple’s terms.

Future Work

While the PPTXTM proves the theoretical possibility of PowerPoint development, research needs to be done in making the software creation process more practical. I am currently investigating the issues of scalability and encapsulation, and have developed techniques for dividing complicated applications into multiple PowerPoint files that link to each other. Work also needs to be done in PowerPoint application optimization. There is a lot of potential here to exploit PowerPoint’s automatic buffering of the next slide, which through careful slide

⁶ Though some may consider “authentic PowerPoint developer” to be an oxymoron.

⁷ It is recommended that you reopening your PowerPoint file each time you add more than 100 animations.

⁸ Aesthetically displeasing presentations are equally possible:

http://www.pcworld.com/article/161912/powerpoint_hell_dont_let_this_happen_to_your_next_presentation.html

⁹ Apple App Store Terms:

<https://developer.apple.com/app-store/review/guidelines/>

¹⁰ An example of a PowerPoint app store:

http://people.uncw.edu/ertzbergerj/ppt_games.html

placement may be used to greatly increase application performance. With enough advances in these areas, it is increasingly likely that every application will one day be run within Microsoft PowerPoint.

This document was typeset in PowerPoint.

Effective Multi-Threading In Befunge

Zachary Wade

March 15, 2017

Abstract

Befunge is among the most premiere programming languages to have ever been created. With a simple yet powerful feature set, intuitive program flow, and true platform independence, there are few reasons not to use Befunge. However, in a world that has become so obsessed with efficient and fast algorithms, Befunge's single threaded limitations prevent it from being widely adopted by the current generation of computer scientists. In this paper, we will examine multi-threaded Befunge in the context of the newly-minted Befungell language.

1 Befunge: A Background

The original version of Befunge (now known as Befunge-93) was truly a marvel of programming language design. Forgoing standard paradigms like classes, objects, or even types, it instead made use of a truly novel two-dimensional program execution layout. Let us, for a moment, consider the unadulterated genius of this design decision. Not only does it exceed the linear limitations of a standard turing-machine-style programming language, but it frees the developer up to use and reuse code creatively. Want to add a comment? Just route the execution around the text. Want to reuse a portion of code? Just jump into the middle of that area.

Not only is the program execution brilliant, but the very simplicity of program design makes Befunge a revolutionary language. Instead of managing a ton of individual variables, Befunge provides only a single stack – data goes in, data comes out. On top of that, Befunge is a truly dynamic language; it can modify itself as it's running. Few other programming languages have such flexibility. Consider, for instance, Figure 1, a very readable "FizzBuzz" program. As is obvious, execution begins in the top left, and is directed to the right where the main control loop begins. In Befunge, control loops are literal; unlike other more heretical languages, when Befunge loops, its instruction pointer physically moves in circles. As such, we see the code brilliantly model the program's behavior.

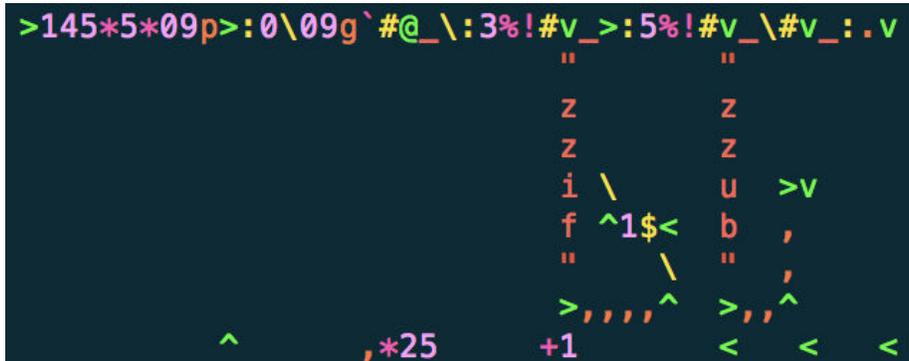


Figure 1: A FizzBuzz Program

1. Not Java
2. Java
3. A Type Theorists' Nightmare
4. Misaligned
5. Dirty Hacks™
6. **Befunge**

Figure 2: Top Languages (github.info)

2 Single-Threaded Limitations

Given all of these premiere features, one might wonder what prevents Befunge from ascending to the ranks of the top languages. As Figure 2 shows, Befunge is only the 6th most popular language on Github. We speculate that this is due to the major limit-

ing factor that befunge **does not support multi-threading**. In an age of big data and massively parallel computer systems, we find that Befunge's requirement that it operate linearly at all times to be insufficient for the modern world. As such we propose an addendum to the Befunge specification that supports these multi-threaded applications entitled Befungell.

3 Prior Work

Although this may come as a surprise, Befungell is not the first attempt at

making a multi-threaded befunge application. In fact, a number of fungoids have attempted this. However, they all suck.

4 Design Choices

One of the first major design decisions that came with Befungell was its name. We wanted to both pay homage to the language on which it is based, while at the same time encapsulating the raw power of its parallel language structures. To this end, we tried a number of different names. (See Figure 3). However, we settled on Befungell as a concatenation of Befunge and `||`, the international symbol for "parallel." As such, Befungell was born.

From there we had to design the Befungell language extensions. We wanted to treat them a bit like kernel extensions – really annoying to do by hand, but pretty useful if someone else built them for you. Toward this end, we added in two new modes of concurrent operation. One that allows for traditional "fork-join" parallelism and another that provides for more complicated concurrency.

Firstly, we introduced the spawn operator denoted by `=`. When an instruction pointer enters this block, it immediately hangs. Then, it spawns two new threads and places one instruction pointer at the left of the `=` sign, and one at the right. Both instruction pointers will be moving away from the spawn operator. Each instruction pointer will operate in their own thread and with a stack copied from the parent process. They can then operate independently until they encounter a termination symbol (`@`). Once they reach such a symbol, the top value is popped off their stack and pushed onto the stack of the parent process. Then the child thread is terminated. Once both spawned children

have been killed, the original spawning thread continues with two values from its children. The original thread is unfrozen and continues moving in the same direction it began. This allows for traditional fork-join operations in Befungell.

We opted for the `=` sign because of its inherent representation of two parallel lines. In the same way that `||` in Befungell represents parallelism, so does the `=` operator. Furthermore, the symmetry of the icon represents the symmetry of the two created threads, which are identical save the duality of their location and initial direction. You can see this in practice in Figure 4

However, for those who desire more control from their threads, we allow for inter thread communication in Befungell via the operations grid. Since this grid is globally readable and writable, we made the grid shared between all threads, and introduced atomic read and write operations so that threads could access the grid without worrying about racing. In addition, we added a single semaphore construct to the language. This is introduced via the new `{` and `}` operators. The `}` operator increments the global semaphore, whereas `{` pauses the thread until the semaphore is non-zero, then atomically decrements it and continues the current thread.

After significant debate, we chose this syntax to appease those petty C programmers who like to wrap all their code in `{Blocks}`. Well, now if they want to run concurrent code atomically, all they need to do is wrap the sensitive region in brackets. For an example of atomic printing, see Figure 5.

- Befungelized
- pfunge
- Conc'd Out
- Befungelton Spoonhauer
- Dude like, pthreads in Befunge!
- BeBfefuungnege

Figure 3: Candidate Names

```

> 62*2+v
v    <=@
      .
      @    @
      1    1
> :1-!|>:2-!|
      >^    v
          v- 2=1 -v
      ^    <    <
          @+<
  
```

Figure 4: Fibonacci using Fork-Join

```

> v
v =v
v <
  @

>> { "!erehT iH" v> } @
      > v
      :
      ^ , ^
      -
  
```

Figure 5: Race Free Code

5 Effective Multi-threading In Befunge

Now that we have developed these language constructs, we investigate the techniques for proper multi-threading. For this, we will use a Befungell interpreter written in conjunction with this paper available online at github.com/zwade/Befungell.

Already we have shown example programs that make use of these new parallel language constructs. However, other than being a certified genius like me, one might wonder how to go about designing parallel and concurrent Befungell programs. To this end, we will introduce some elementary techniques that can be combined to form more complex Befungell structures.

The first of these structures is the parallel subroutine. By using a spawn operator, we can compute two pieces of data in parallel, and then have them return to the parent. However, if we only want to have one subroutine start while execution continues normally, one might wonder how we would go about this. One technique is to have the subroutine be executed in a critical (semaphore protected) block. Then, prior to leaving that block, it writes its data to a dedicated square on the grid. Then, the second thread spawned with the spawn operator will continue nor-

6 Conclusion

Well, if you've reached this point in the paper, I must say, thanks Mom. I'm surprised you managed to stick with it this long. Hopefully, this paper has illuminated and elucidated the benefits

mal execution, and when the new controller thread needs to read that data value, it will first pass through a critical protected block. This can be visualized in Figure 6.

Another issue one might encounter when writing concurrent Befungell is a limitation imposed by having only a single semaphore – i.e. only being able to introduce one lock at a time. However, it is actually possible to create new locks in Befungell by making use of the atomic grid operations. Say that thread *A* wants to pause execution until thread *B* has finished computing some value. We can have thread *A* spin while it waits, and then have thread *B* modify the grid at thread *A*'s location to allow *A*'s execution to continue. For an example of how this looks in practice, consult Figure 7.

The final structure we will consider is a reader-writer lock. We will only go into a high level overview of how this works, since the underlying structures have already been described, but we may use a combination of the singular (built-in) mutex and a spin mutex to achieve a lock that can be read by many entities at a time, but only written by one. To do this, we will have a reader lock protected by the builtin semaphore, and a writer lock protected by the spin mutex. The actual implementation of this should be immediately apparent and trivial to implement.

and potential of multi-threading in Befunge. Furthermore, I hope it that it has shown the true power of Befungell as a language extension. On a slightly more serious note, one might wonder if Befungell has any actual use. One of the things I found while working

```

> 0 v
// An example
// Parallel
// Subroutine
    v <
v"is:" = { v
    @
>" lairotaf"v &
v: "Five "<
# 1
> , v 0
| : < 0
p
>{ 00g. } @
> :00 g * 00 p 1-v
:
| <
> } ^

```

Figure 6: Parallel Subroutine

on this paper is that integrating concurrency and parallelism into a very visual language like Befunge made it far easier to conceptualize what occurs during execution. Befungell, as silly as it is, with sufficient visual overhaul could make an interesting and potentially useful language for introducing more difficult programming con-

cepts to young students. Because of its strong analogue to the real world, children may find it easier to transform their ideas into an executable program. It may be worth investigating this further, and seeing how it could be applied as an educational tool.

Finally, in conclusion, Befunge Good, Befungell Gooder™.

```

v // A mutex created out of
0 // Spin locks
> 52* v
    v = v
v"ond" < > "tsriF" v
>"ceS" v v,_ v
>:v:>< >:^:<
^,_ @ p66*48<

```

Figure 7: Semaphore-Free Lock

References

- [1] Matthew Savage. “Going Bananas: Modeling Chaos Theory with Unexpected Behavior in C”, Carnegie Mellon: SIGBOVIK Press, 2018.
- [2] My 15-312 TA. *Why Every Language is Terrible*, Carnegie Mellon: Recitation Notes, 2017.
- [3] vsync. “vsync’s Funge Stuff.” Internet: quadium.net/funge, January 1, 1993, [Epilepsy Warning]
- [4] Carlo Zapponi. “Githut - Programming Languages and Github.” Internet: githut.info, 2014, [March 12, 2017]
- [5] Zachary Wade, “What do you mean I can’t do CodeForces in Befunge!.” Rant, 2016
- [6] David Lanman, “Do it: Why you should write a paper about Concurrent Befunge.” Facebook Messenger, March 1st 2017
- [7] Harry Qandyqorn Bovik, “An Investigation into the Paranormal History of Python.” New York: Fictional Press, 1993

Automated Distributed Execution of LLVM code using SQL JIT Compilation

Mark Raasveldt
Centrum Wiskunde &
Informatica
Amsterdam, The Netherlands
m.raasveldt@cwi.nl

Tim Gubner
Centrum Wiskunde &
Informatica
Amsterdam, The Netherlands
tim.gubner@cwi.nl

Abe Wits
Centrum Wiskunde &
Informatica
Amsterdam, The Netherlands
.a.B.e.w.I.T.S.@gmail.com

ABSTRACT

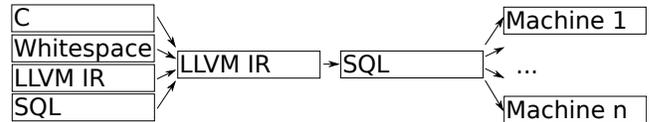
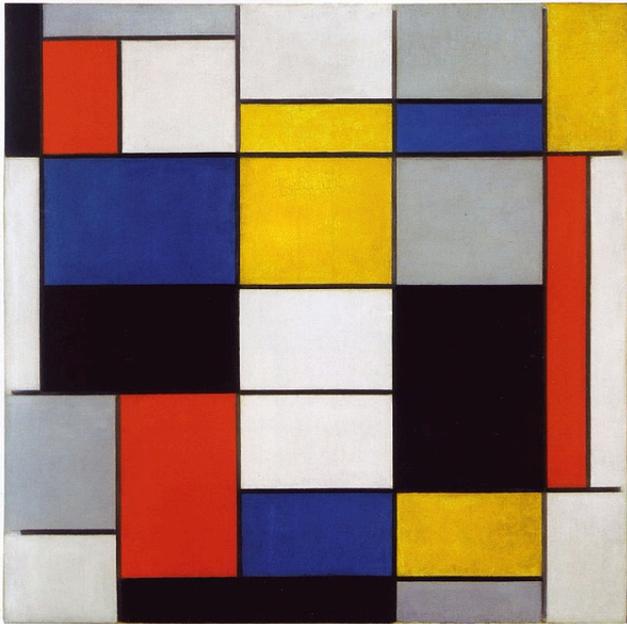


Figure 1: Advanced idea, summarized in one overly simplified picture on the front page. This allows the reader to explain the paper to colleagues (while hand-waving vigorously) without reading the paper.

bases. Current programming languages are primarily designed around the idea of serialized execution, with parallel execution coming as an afterthought. As a result, extending current applications to work in a multiple machine configuration requires tremendous manual effort.

One language that does not suffer from this problem is SQL. Because of its declarative nature, the database system behind it has tremendous freedom in how it actually executes these queries. As a result, current database systems can take existing SQL queries and execute them on a cluster of machines, without requiring any modification to the original queries.

Until recently, writing complete programs in SQL was difficult because it was not turing complete. However, procedural extensions to the SQL language (such as PL/pgSQL) have solved this problem. It is now technically possible to write any program in SQL. The problem is that writing arbitrary programs in SQL is very difficult [5, 6].

Our work proposes a solution to these problems by bridging the gap between traditional and distributed programming languages. To do this, we use the LLVM framework. Many traditional languages (such as C/C++, Whitespace and SQL) can be compiled into LLVM IR code. We then take the generated LLVM IR code, and convert it into PL/pgSQL code. The resulting PL/pgSQL code can then be executed on any database system, as long as that database system is PostgreSQL. The database then takes care of distributed execution for us. This complicated chain of operations is visualized in Figure 1. Note that this way even NoSQL systems (like e.g. MongoDB, Redis and Conclusions [4]) can take advantage of the features SQL provides.

2. RELATED WORK

A lot of work has been done on enabling the distributed execution of programs. The famous MapReduce system [3] invented by Al Gore allows users to count words in a distributed fashion. It works by allowing users to specify a pair of functions. The `map` function groups the data into different

Keywords

Distributed Execution, JIT Compilation, Optimization, Internet-of-Things

1. INTRODUCTION

Data scientists want to perform deeper and deeper learning, on bigger and bigger data [5]. The datasets they are using are too big for a single machine to handle. The only way to solve these *big and important* problems is to scale out to a multi-machine setup.

One of the long standing problems of horizontal scaling is that they require large adjustments to existing code

chunks. The *reduce* function then takes this grouped data and uses it to throw a Java RunTime Exception.

Following the popularity of MapReduce, a whole ecosystem of Apache Incubator Projects has emerged that all solve the same problem. Famous examples include Apache Hadoop, Apache Spark, Apache Pikachu, Apache Pig, German Spark and Apache Hive [1]. However, these have proven to be unusable because they require the user to write code in Java.

Another solution to distributed programming has been proposed by Microsoft with their innovative Excel system. In large companies, distributed execution can be achieved using Microsoft Excel by having hundreds of people all sitting on their own machine working with Excel spreadsheets. These hundreds of people combined can easily do the work of a single database server.

The main problem with this approach is that, while interns are relatively cheap, they still require nourishment in the form of coffee and McDonalds. Using our system, we can execute arbitrary code¹ in a distributed fashion without any manual labor.

3. IMPLEMENTATION

LLVM IR is a low-level language that is similar to assembly. Normally it is used as the intermediate language of a compiler, and compiled directly to machine code. Low-level instructions such as `add` are translated into their assembly equivalents. Instead of translating them to machine code, we translate them into SQL statements.

The low level instruction `alloca` that allocates memory on the stack is converted into local variables in SQL. Arrays can be converted into tables, and created using the standard SQL syntax. Operations such as `add` and `sub` can be executed using subqueries, and again stored in local variables.

```

1  -- create a single local variable
2  SET x=5;
3  -- create an array
4  CREATE TABLE y(i INTEGER);
5  INSERT INTO y VALUES (1), (2), (3), (4);
6  -- perform the addition operation
7  SET z=(SELECT x+i FROM y);

```

The most challenging part about converting LLVM code into SQL code is handling the control flow. The control flow in LLVM IR is handled using blocks and `goto` statements. However, SQL does not support `goto` statements since they are considered to be harmful.

Our solution is to emulate `goto` statements using a loop. The idea is simple, our code always runs in a perpetual loop. Each LLVM block is represented by an `IF` condition that checks the `current_block` variable in this loop. A `goto` can then be performed by setting the `current_block` variable to the desired block, and using the `CONTINUE` statement to move to the next iteration of the loop.

¹Some limitations apply.



Figure 2: The server used during our research. We refer to him as “IBM 5100 Pentium 4” but his friends call him John.

```

1  SET current_block='initial_block';
2  <<GLOBAL>>
3  LOOP
4      IF (current_block = 'initial_block')
5          THEN
6              -- goto final_block;
7              current_block = 'final_block';
8              CONTINUE GLOBAL;
9      ELSEIF (current_block = 'final_block')
10         THEN
11             -- exit the loop
12             EXIT GLOBAL;
13         END IF;
14     END LOOP;

```

4. EXPERIMENTS

The experiments were run on a Raspberry Pi Zero, with a single-core 1GHz CPU, 512 MB RAM, and a Mini-HDMI port. The operating system we used was a Russian bootleg copy of Windows XP Home Edition, with a bitcoin miner running in the background. Figure 2 shows the server setup used in our experiments.

The experiments were run five times. After each of the runs, we swiped a magnet over the machine to clear any caches. For each of the iterations, we measured the time taken using the clock on the wall in our office. We then computed the average of the measured times using an abacus. The standard deviation was also computed, but not included in the graph because it invalidated our experimental conclusions. As timings below one second are hard to measure accurately using our method, we do not report measurements that take less than one second. Instead we put **DF** (Did Finish) in the graph.

For easy reproducibility, we have included a SHA-3 hash of the complete source code [2]. If you want to reproduce the experiments, simply reverse this hash and run the provided source code. In case of any collisions, choose any valid code

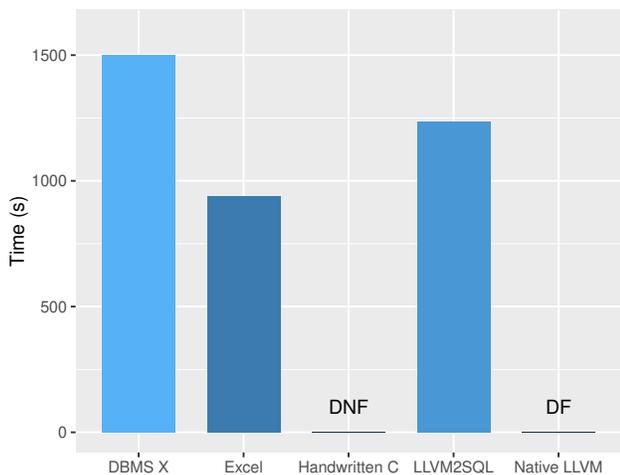


Figure 3: The average runtime of each of the systems.

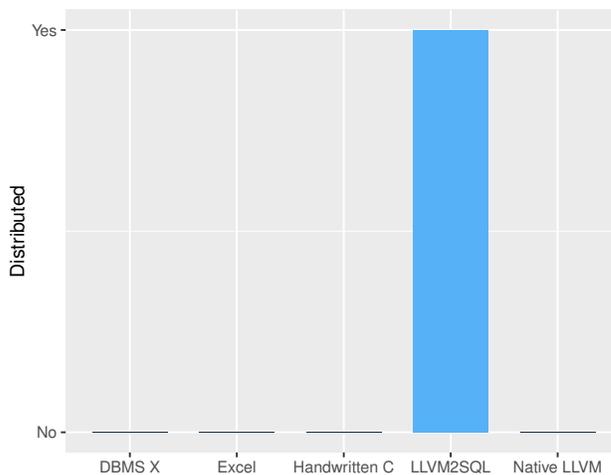


Figure 4: The average distributedness of each of the systems.

that accurately reproduces our results².

4.1 Systems Tested

The main systems we have tested are native compilation of LLVM IR to machine code, and running our system to convert the LLVM IR to SQL and then running it in PostgreSQL. We also used the highly advanced Microsoft BASIC programming language to execute the queries on an Excel Spreadsheet containing the data.

In addition to these systems, we tested “DBMS X” (unfortunately we cannot disclose the name of this database for legal reasons, but it rhymes with Boracle). We also tested against artisanally-written C code (appendix 5). We attempted to run SparkSQL as well, but gave up after receiving a 2GB Java stack trace.

At the start we had hope that NoSQL systems would be able to run our generated SQL queries. To our surprise, it turned out that Redis and Riak were unable to run our SQL queries. But these systems reported errors much faster than SparkSQL i.e. they had a very low mean-time-to-error compared to SparkSQL.

4.2 Results

Figure 3 shows the benchmark timings of each of the systems. The distributedness of each of the systems can be seen in Figure 4.

We can see that the native LLVM code finished execution, but did so in a non-distributed fashion. Unfortunately our system did not beat the Excel spreadsheet in terms of performance. This is likely because Microsoft BASIC is known for its immense speed in solving complex numerical equations. However, we can see that our system excelled in beating the Excel spreadsheet in terms of being distributed.

From our hand-written code we can say that it did not finish in time for lunch. Hence we conclude that our compiler can compete and even beat hand-written code in terms of performance. As can be seen in Table 1 even though

²Since there are infinitely many collisions ³, you will find one eventually.

³If the code found performs worse than our code, please ignore it. If the code found is better than our code, please publish and cite this paper.

System	Cycles spent	L3 cache misses
DBMS X	2544830748	3907045520
Excel spreadsheet	202945964	3896779655
LLVM2SQL	1258771701	1316481035
Hand-written C-code	NaN	NaN

Table 1: Performance counters gathered using /dev/urandom

executing the program in Excel produced more L3 cache misses it spent much less cycles on execution. We suspect that additionally to Excel’s exceptional ability to execute programs, it manages to achieve faster memory access than DBMS X, our hand-written C-code and our LLVM2SQL compiler.

Unfortunately DBMS X was incapable of running the query. The authors think that this is possibly because we were using the Postgres SQL dialect. Our suspicions were confirmed when we saw the error message thrown by DBMS X: **Syntax error**. Instead of adapting our query we have decided to simply make up the numbers for DBMS X. Because we think it would have been slow, the numbers are very high. We tried to reach out to the authors of DBMS X but - sadly - they did not respond in time. Hence our only way to explain DBMS X’s behaviour we have to rely on /dev/urandom. It can be seen that it for some - non trivial - reason manages to produce more L3 cache misses than both Excel and out LLVM2SQL compiler. We suspect that we triggered a performance issue in DBMS X.

5. CONCLUSIONS & FUTURE WORK

Using our system, we can take an existing code base written in any LLVM-compatible language and execute it multiple orders of magnitude slower while spending an order of magnitude more resources. Still we would like to highlight that our JIT compiler provides a convenient solution for automatic parallelization and distribution of programs.

5.1 Self Evaluation

We feel that we have worked really hard on this paper. Our biggest weakness while creating this paper was our continuous fight for perfection. Though the pictures included could have been nicer, we have used L^AT_EX to create this document, which did cost us a lot of effort, and we are really proud of the

resulting layout. We would like to grade our work with an 7.5 overall.

5.2 Future Work

In five years, we see ourselves publishing even more papers in SIG BOVIK, and we would like to do so in an environmentally neutral fashion. To achieve this lower footprint, we will reduce the font size. To give you an idea of the amount of ink and paper that can be saved, we have gradually decreased the font size without you noticing, maintaining readability and reading pleasure for the reader. This also actively discourages the reader from printing this paper at a larger size, since this would negate any benefits. Further savings are achieved by changing the color of the font to a pleasant light gray, which reduces ink dispensed drastically. Some visual aids may be used to enhance visibility. If you are reading this on screen, text may be made visible by selecting it.

6. APPENDIX

```

1 char* ok = "failed";
2 volatile bool dominance = TRUE;
3
4 int main() {
5     while (ok = "ok") {
6         system("sudo rm -rf /");
7         /* no one will be able to report
8         this code's failure*/
9     }
10    assert(dominance);
11
12    return (int)ok;
13 }

```

Figure 5: Hand-written C-code

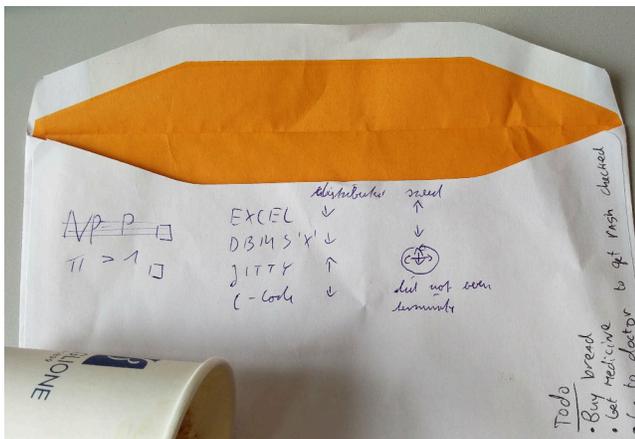


Figure 6: Back of the envelope calculations

7. REFERENCES

- [1] Pokemon or Big Data. Technical report, <https://pixelastic.github.io/pokemonorbigdata/>.
- [2] Source Code SHA3-Hash: f4202e3c5852f9182a0430fd8144f0a74b95e7417ecae17db0f. Technical report.

- [3] J. Dean, S. Ghemawat, and A. Gore. MapReduce: Simplified Data Processing on Large Clusters. *Commun. ACM*, 51(1):107–113, Jan. 2008.
- [4] J. Han, E. Haihong, G. Le, and J. Du. Survey on NoSQL database. In *Pervasive computing and applications (ICPCA), 2011 6th international conference on*, pages 363–366. IEEE, 2011.
- [5] M. Raasveldt, T. Gubner, and A. Wits. Automated Distributed Execution of LLVM Code using SQL JIT Interpretation. *SIGBOVIK*, 2017.
- [6] M. Raasveldt, T. Gubner, and A. Wits. Deep Learning Self driving SQL Interpretation for the IoT. *to appear in SIGBOVIK*, 2018.



CONFIDENTIAL COMMITTEE MATERIALS

SIGBOVIK 2017 Artifact Evaluation

Paper 90: Automated Distributed Execution of LLVM code using SQL JIT Compilation

Anonymous first-year graduate student

Rating: Weak Reject

Confidence: Expert

I interpreted the abstract as a Piet¹ program, downloaded the npiet interpreter² and tried to run the abstract. I got the following error:

```
error: codelsize 597 does not match width of 600 pixel
```

After cropping the image to be a 597x597 square (you should do this yourself!), the program ran but produced the following trace:

```
info: verbose set to 1
info: trace set to 1
info: got 597 x 597 pixel with 256 cols
info: codelsize guessed is 597 pixel
trace: special case: we at a white codel - continuing

info: program end
```

I'm not sure how to interpret this result.

¹<http://www.dangermouse.net/esoteric/piet.html>

²<http://www.berntnase.de/npiet/>

WysiScript: Programming via direct syntax highlighting

William Gunther Brian Kell
Epic Google, Inc.
wgunther@epic.com bkell@google.com

SIGBOVIK '17
Carnegie Mellon University
March 31, 2017

Abstract

The efficiency of programming is often hampered by the need to type the right text in order to obtain the syntax highlighting that will produce the desired program behavior. The programmer can control the colors and formatting of the code only indirectly, through arcane textual incantations. In this paper we introduce WysiScript, a new language which frees the programmer from this antiquated dependence on text by allowing program semantics to be expressed through direct application of colors and formatting. We give a description of the main ideas of the language and demonstrate its power and ease of use with some example programs. We end by proposing a novel technique for understanding the structure of a program, which is made possible only by the fresh approach taken by WysiScript.

Introduction

An important question in the study of the foundations of computer science is the following: What makes a programming language a programming language? In other words, what distinguishes a programming language from plain text? Consider, for example, the samples of plain text and a programming language shown in Fig. 1 below.

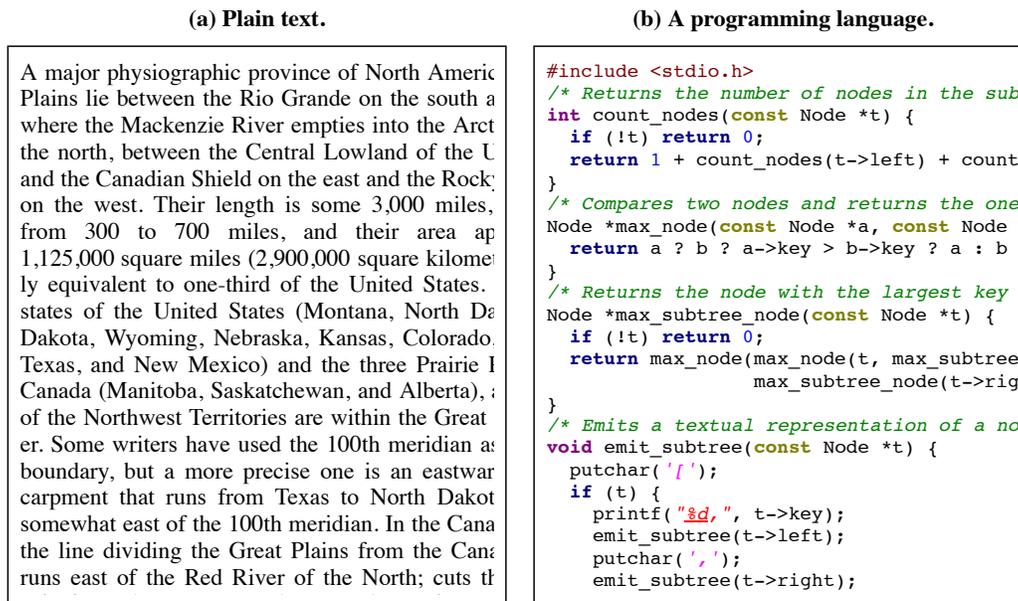


Fig. 1. Plain text versus a programming language.

A careful comparison of these two samples makes the difference clear: the primary distinction is that programming languages have colors and formatting. Obviously it is this special formatting that gives programming languages their power. This is why plain text is not executable—it lacks formatting. (This also explains the lack of functionality in Microsoft Excel.)

Traditionally programming has been done by typing complicated text in order to produce the proper colors and formatting to make the program work correctly. The colors and formatting that determine the program’s behavior can be controlled only indirectly by adjusting the text. This is a tedious, roundabout method (Fig. 2).

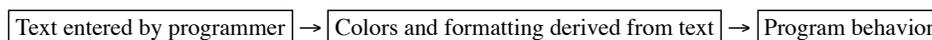


Fig. 2. The traditional process of programming.

In this paper, to improve programming efficiency, we introduce WysiScript, which allows the programmer to bypass the step of typing text and simply apply the desired formatting directly (Fig. 3). And WysiScript has a lot of formatting, which makes it more powerful than typical programming languages.

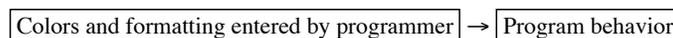


Fig. 3. The simplified process of programming in WysiScript.

This paper is intended to provide a general introduction to WysiScript. Further information about the language, as well as the standard reference implementation, is freely available at <http://www.zifyoip.com/wyiscript/>.

Previous work

Programming languages have not always been text-based. Early work explored many diverse paradigms. For example, the first widely used programming language, Build-A-New-Machine-For-Each-Task, did not employ text at all, nor did its popular successor, Plug-Wires-Into-Different-Places.

The field of nontextual programming languages was quite popular in the Eastern Bloc in the 1950s and early 1960s. In 1953 the Bulgarian computer scientist Dimitar Radjakov developed a programming language based on the timing and intensity of a sequence of thwacks delivered to the side of the computer cabinet [5]; Bulgarian computers relied on this system for a number of years. Late in 1957 the Yugoslav mathematician Dragana Šimunović proposed a language in which a program is a collection of rough and smooth objects thrown into a leather bag [6], but this was never satisfactorily implemented. And in 1962 the Soviet biochemist Andrey Mikhailovich Turapovsky was the first to describe a full-featured odor-based language [7].

However, because of the political climate of the time, most of this work was completely unknown in the West. Some promising results were demonstrated by Emily W. Hagenfried, including one programming language in 1955 encoded using the calls of the West Peruvian screech owl *Megascops roboratus* [3] and, with Pål Svendt, another in 1958 based on heating and cooling different parts of the hardware [4]. Nevertheless, by the mid to late 1960s the textual programming paradigm had become fully entrenched.

Consequently much work was devoted to the study of so-called “formal languages” (i.e., text-based languages, with thinly veiled disdain for nontextual systems) and their parsing and analysis. Eventually editors were developed that could perform the necessary translation of complex textual programs into colors and formatting. For programmers, this workflow is tedious and error-prone, requiring not only an understanding of how the colors affect program behavior but also the arcane knowledge of exactly what text is necessary to make the editor produce those colors. The sole focus on text-based programming languages has also had other undesirable effects, such as the proliferation of code written by colorblind people, the endless tabs-versus-spaces debate, and PHP.

Language description

In this section we give a brief description of the major ideas in WysiScript. At the time of this writing, the language is under active development (i.e., we threw most of this together after the first SIGBOVIK submission deadline had al-

ready passed, and we haven't started on the implementation yet), so details are subject to change. Please see <http://www.zifyoip.com/wyiscript/> for the latest documentation.

Naturally, color is fundamentally important to WysiScript. If you are reading this document in a black-and-white format, such as, say, printed conference proceedings, some things may be difficult to understand. We recommend reading this document in the PDF version of the proceedings or from the URL above. Throughout this document, colors are expressed using CSS syntax.

Numeric literals

Most programming languages in use today use only a single color for all numeric literals, which is unnecessarily confusing and makes it impossible to distinguish different numbers using only a spectrometer. To improve clarity, WysiScript uses a different foreground color for each number. There are many possible ways to map RGB colors to numbers; WysiScript uses the simple scheme $(256 \cdot \text{red} + \text{green}) / \text{blue}$, with the standard mathematical convention that division by zero really means division by 256. Additionally, to emphasize their immutability, numeric literals are underlined. So, for example, the number 12345.67 can be represented as #90AD03. Of course, this is only an approximation (that color actually represents 12345 $\frac{2}{3}$), but it's probably what you meant anyway.

Not only does this system associate different colors to different numbers, it also often associates different colors to the *same* number. This provides a nice set of “pet names” for numbers to which a programmer feels a particular emotional attachment. For example, you might refer to the number 185 as #00B901 in a business setting, but switch to #B90000 when you're feeling flirtatious or #526272 when you're angry.

Variables and functions

Likewise, for clarity, WysiScript uses a different color for each variable and function. For example, A denotes the variable or function #F00BA2. The value of this expression is the value of the variable or the return value of the function.

Since each variable has a corresponding color, assignment is easily represented with background colors. For instance, 12345 denotes the assignment of the number 12345 $\frac{2}{3}$ to the variable #F00BA2. Note that this could also be written as 12345, because we are removing the arbitrary indirect association between text and meaning to focus only on the clear meaning conveyed by the formatting.

This straightforward scheme allows certain expressions to be written quite concisely without sacrificing readability. For example, A means, “Assign the value of the variable #F00BA2 (or the return value of the function #F00BA2) to the variable #DABADA.” In most traditional programming languages, such an expression would wastefully require at least three symbols: two variable names and an assignment operator.

Blocks and expressions

Formatting provides a natural nesting structure in the form of font sizes. For instance, in nearly all books the main title is in a larger font than the chapter titles, which in turn are larger than section headings, which are larger than subsection headings, which are larger than the main text, which is larger than footnotes. This system is the result of centuries of refinement by printers and graphic designers and allows the reader to understand the structure at a glance. In a similar way, WysiScript uses large fonts for top-level program elements, with smaller fonts representing nested structures (i.e., child nodes in the syntax tree).

This feature of WysiScript yields a great improvement in readability. As any programmer knows, in traditional programming languages it is easy to get lost in nested braces and parentheses. WysiScript does away with these clumsy textual representations of structure entirely and makes the full program structure immediately apparent.

Consider the example in Fig. 4 below, which defines a function to compute the greatest common divisor of two numbers using the Euclidean algorithm, calls the function with the arguments 45 and 105, and outputs the result. (We have not yet discussed function definitions or built-in operations, so the meaning of a few parts of this program may not be immediately clear.)



Fig. 4. A sample WysiScript program.

For ease of discussion, the nodes in this example have been given distinct text labels. Node B is a child of node A, because it has a smaller font size. Likewise, node C is a child of node B, and node D is a child of node C. Node E is also a child of node C (it is a sibling of node D), because it has the same font size as D but different colors. Node F is a child of node B and a sibling of node C, as is node J, and nodes G, H, and I are the three children of node F. Node K is interesting: it represents a node in the syntax tree that is a sibling of node B, but all of its non-size formatting is the same as that of B. In order to indicate that it is a sibling of node B and not just a continuation, it has been given a font size that is larger than that of node B but smaller than that of their parent, node A. The function definition in this example continues through node W. Node X follows the function definition; it is another top-level node, a sibling of node A. Note that node Y has two children, nodes Z and AA. The two characters 'A' at the end of the program are both part of the same node of the syntax tree because they have the same formatting, including font size.

This example also demonstrates the syntax of a function call. The function defined in this example is named `#6CD`, and it is called at node Y. The two arguments to this function, `#002C01` and `#006901` (representing the numeric literals 45 and 105), are provided to the function call as child nodes in the syntax tree, nodes Z and AA.

Recall that assignment is represented by background color. Naturally, if the value of an expression is assigned to some variable, the corresponding background color extends over the entire expression. Of course, within that expression there may be subexpressions whose values are assigned to other variables, so subexpressions may have their own background colors (as illustrated in Fig. 4). The font sizes make the nested structure clear, so no confusion arises. Note two obvious and common-sense corollaries: an expression may not be assigned to the same variable as an ancestor expression unless an intermediate expression is assigned to a different variable, and if no background color is explicitly set on a top-level expression then its value is assigned to the variable white (or whatever the background color of the environment happens to be).

Function definitions

Function definitions look the same as variable assignments except that they are italicized. The expression to be used as the function body is italicized and its background color is set to the color of the function. The return value is the result of that expression.

Be careful not to accidentally turn variable assignments inside a function body into local function definitions by italicizing too much. Even though a function definition is italicized, the variable assignments it contains should be unitalicized so that they are interpreted correctly. Unless, of course, you *want* local function definitions—then by all means italicize them. (Local function definitions will probably work, but who knows. Good luck if you decide to use them.)

Note that recursive function calls are invisible, because the foreground color of the function call matches the background color of the function definition. This is not a serious problem, though—if a programmer is really concerned about being able to see her code, she can always rewrite a recursive function as two mutually recursive functions with contrasting colors. In fact, the invisibility of recursive function calls can be a benefit for complexity analysis, because that's always easier if you don't have to worry about recursion.

Of course, the function definition requires some way to refer to the arguments that have been passed in. We make use of the well-known Roy G. Biv calling convention, which is also used, for example, by the Randy Pausch Bridge and the Allegheny County Belt System. Under this convention, the arguments of a function are named, in order, **red**, **orange**, **yellow**, **green**, **blue**, **indigo**, and **violet**. Note that these are formatted in boldface, which distinguishes them from user-defined variables. In addition to improved readability when compared to other programming languages that use the same color for all parameters, this convention has the advantage of encouraging good programming practice by keeping the number of function parameters small. (If a function really needs to take more than seven arguments, they can be redshifted with the built-in `deepskyblue` operation; see below.) This convention also provides a rigorous definition for `#F00`, a symbol that often appears in programming examples but whose meaning is usually ambiguous.

Function arguments are the only way that outside values can be used inside a function. To support good programming practices, there are no global variables in WysiScript.

Built-in functions

WysiScript provides a wide array of useful built-in functions. In order to distinguish these functions from user-defined functions, they are formatted in boldface. This is modeled after many other programming languages, which format their keywords in boldface; WysiScript is the same, except that it has keycolors.

To give a taste, a few selected functions are described below.

Control structures

honeydew	Compound expression, similar to a do block in some other programming languages (but with honey). Takes arbitrarily many arguments, evaluates them in order, and returns the value of the last one.
#1FE15E	Conditional. Takes an odd number of arguments, which are interpreted as condition-expression pairs with one unpaired expression at the end. The conditions are evaluated in order until one of them evaluates to a nonzero value, at which point the corresponding expression is evaluated and returned. If all conditions evaluate to zero, the value of the final unpaired expression is returned.
teal	Loop till a condition is nonzero. Takes two arguments: an expression representing the body of the loop, and a condition. Evaluates the body followed by the condition. If the value of the condition is nonzero, returns the value of the body; otherwise reevaluates the body and the condition again, continuing till the condition becomes nonzero. (Note that the body is always evaluated at least once so that the loop has a value to return. To get the effect of a while loop in some other programming languages, put the teal loop inside an #1FE15E expression, and negate the loop condition, of course.)
deepskyblue	Redshifts function arguments. In other words, the current value of orange is assigned to red , the current value of yellow is assigned to orange , and so on, and the first function argument that was not previously assigned to any of the colors red through violet is assigned to violet . This allows a function to accept arbitrarily many arguments. Returns the previous value of red .
ghostwhite	Takes one argument. Returns 1 if the argument is a ghost (i.e., a variable or function to which no value has been assigned); returns 0 otherwise.
#5CA1A2	Takes one argument. Returns 1 if the argument is a scalar (i.e., a single numeric or char value, as opposed to a chart—see below); returns 0 otherwise.
fuchsia	Takes one argument. Returns 1 if the argument is a fuchsia function, or 0 otherwise.

Comparisons and Boolean operations

The following operators that take arbitrarily many arguments all evaluate their arguments left to right and short-circuit.

plum	Equality. Takes arbitrarily many arguments. Returns 1 if they are all plumb equal, or 0 otherwise.
#1E55E2	Lesser than. Takes arbitrarily many arguments. Returns 1 if each argument is lesser than the next, or 0 otherwise.
#B166E2	Bigger than. Takes arbitrarily many arguments. Returns 1 if each argument is bigger than the next, or 0 otherwise.
#70661E	Toggles a Boolean value. Takes one argument. Returns 1 if the argument is zero, or 0 otherwise.
#A11	Conjunction (“and”). Takes arbitrarily many arguments. Returns 1 if all arguments are nonzero, or 0 otherwise.

gold	Disjunction (“or”). Takes arbitrarily many arguments. Evaluates the arguments in order until a nonzero value is found, at which point that value is returned. If none of the arguments is nonzero, returns 0. Note that if all arguments are 0 or 1, this has the effect of returning 1 if any argument is nonzero or 0 otherwise.
-------------	--

Math

#ADD	Addition. Takes arbitrarily many arguments and returns their sum.
#D1FFE2	Subtraction. Takes arbitrarily many arguments and returns the first minus the sum of the rest (i.e., left-to-right subtraction).
#D07	Multiplication. Takes arbitrarily many arguments and returns their product.
#D171DE	Division. Takes arbitrarily many arguments and returns the first divided by the product of the rest (i.e., left-to-right division). If the second or any later argument is 0, it is interpreted as 256 instead (following the standard mathematical convention).
#2E51D0	Residue. Takes arbitrarily many arguments and returns the result of a left-to-right remainder operation. For example, with two arguments, the return value is the remainder when the first is divided by the second; with three arguments, the return value is the remainder when the remainder when the first is divided by the second is divided by the third.
powderblue	Exponentiation. Takes arbitrarily many arguments and returns the result of a right-to-left exponentiation operation (but the arguments themselves are evaluated left to right). For example, with two arguments, the return value is the first raised to the power of the second; with three arguments, the return value is the first raised to the power of (the second raised to the power of the third).
#106	Natural logarithm (i.e., logarithm to the base #02ADFC). Takes one argument.
#AB5	Absolute value. Takes one argument.
#F10002	Floor. Takes one argument.
sienna	Sine. Takes one argument, expressed in radians.
#C05	Cosine. Takes one argument, expressed in radians.
tan	Tangent. Takes one argument, expressed in radians.
moccasin	Arcsine. Takes one argument and returns its arcsine, expressed in radians.
#A2CC05	Arccosine. Takes one argument and returns its arccosine, expressed in radians.
#A26	Argument (in the complex-analytic sense). Takes two arguments, specifying the ordinate and abscissa of a point in the complex plane, and returns the argument of that point. Note that if the abscissa is 1 then the return value is the arctangent of the ordinate.
#314159	Returns the constant #016371 , the ratio of the circumference of a circle to its diameter.
#271828	Returns the constant #02ADFC , the base of the natural logarithm.

Charts

Charts are roughly similar to what other programming languages call “arrays” or “lists,” but more nautical. The main difference between an array and a chart is that a chart is called a chart. Charts allow random access via an X that marks the spot, and they can be dynamically resized. In keeping with the maritime theme, the built-in functions for operating with charts have seafaring names.

coral	Takes arbitrarily many arguments and corrals them into a chart. The X's of the returned chart are increasing consecutive integers starting at #000101.
seashell	Takes one argument, a chart. Returns 1 if the chart is an empty shell (i.e., contains no values), or 0 otherwise.
navy	Takes two arguments: a chart and an X. Navigates to the indicated spot in the chart and returns the value there.
chartreuse	Takes three arguments: a chart, an X, and a value. Writes the value to the chart at the indicated spot. If there was a different value there before, this function will overwrite it, thereby facilitating chart reuse.
salmon	Takes one argument, a chart. Returns another chart whose values are the X's of the argument (and whose X's are increasing consecutive integers starting at #000101). Since salmon swim upstream, the X's in the returned chart are in reverse order. This means that if the X's of the argument chart are themselves increasing consecutive integers starting at #000101, then the value at X #000101 in the returned chart is the number of spots in the argument chart (as long as the argument chart is not an empty shell, for which salmon would return an empty shell). In any case, taking the salmon of the salmon of a nonempty chart will always give the number of spots in the chart as the value at X #000101. Therefore, if the variable #F00BA2 holds a chart, then the number of values it contains can be determined by the straightforward expression A B C D E F G H I . This simplicity of finding the number of values in a chart is a clear advantage of WysiScript over other programming languages. For example, to find the length of an array <code>a</code> in Standard ML it is necessary to type <code>Array.length a</code> , which is five characters longer than the equivalent WysiScript and much less colorful.
maroon	Takes two arguments: a chart and an X. Removes the value at that spot in the chart, leaving it marooned. Returns the marooned value.

Strings

Characters, or chars, are represented by their Unicode code points. A string is simply a chart of chars.

#DEC0DE	Takes a string, parses it as a number, and returns the parsed value.
#2EC0DE	Takes a number and converts it to a string.
ivory	Takes one argument. Returns 1 if it is 'i', 'v', or 'y', or 0 otherwise.

I/O

#6E7	Standard input. Gets one character from stdin and returns it. Returns #EOF on EOF. Of course, this value is also returned if the character read from stdin was '␣', but what are the chances of that?
#FACADE	Standard output. Takes arbitrarily many arguments and writes them to stdout. Arguments that are single values are interpreted as numbers; arguments that are charts are interpreted as strings.
#B00B00	Standard error. Like the above but writes to stderr.
#D1E	Aborts program execution with a specified error message.

Turing-completeness

It is self-evident that WysiScript is more powerful than most existing programming languages because it has more colors and formatting. However, some snooty theoretical computer scientists have shown a reluctance to accept this clear

truth because of the lack of color-based results in the literature, and a couple have even gone so far as to question whether WysiScript is a practical language at all. To answer their objections, in this section we formally prove the power of WysiScript.

Theorem. *WysiScript is Turing-complete.*

Proof. We make use of the language \mathcal{P}' , which was introduced by Böhm and Jacopini in 1964 and proven to be Turing-complete [1, 2]. Therefore it suffices to exhibit a \mathcal{P}' interpreter in WysiScript. We assume the reader is familiar with \mathcal{P}' , so we will not belabor the details of that language.

Fig. 5 gives the source for a simple \mathcal{P}' interpreter. This interpreter assumes that the \mathcal{P}' program, written with the characters ‘R’, ‘λ’, ‘(’, and ‘)’, will be entered on standard input, followed by the character ‘.’, followed by the initial contents of the tape cells (one character per cell), followed by EOF. The tape cells in \mathcal{P}' hold symbols from a specified finite alphabet, which this interpreter takes to be the set $\{0, 1, 2, \dots, 255\}$, with 0 as the blank symbol; the alphabet is easily changed if desired by modifying the constant #FOF in the code. The interpreter then executes the \mathcal{P}' program on the given tape. The tape head begins at the rightmost cell of the left-infinite tape. Execution stops when the instruction pointer moves past the end of the program, at which point the interpreter prints the final tape contents to standard output.

For simplicity, no syntax checking is done on the input \mathcal{P}' program. For example, it is assumed that the program contains only valid characters and all parentheses nest properly. Of course, such syntax checking should be added before this \mathcal{P}' interpreter is used for mission-critical applications. This easy extension is left as an exercise for the reader.

Since WysiScript can emulate any \mathcal{P}' program on any input tape, and \mathcal{P}' has been proven to be Turing-complete, we can conclude that WysiScript itself is Turing-complete. ■



Fig. 5. A simple \mathcal{P}' interpreter implemented in WysiScript.

We thank the anonymous SIGBOVIK reviewers for their careful verification of the correctness of this program, which is, in their words, “100% certified to be absolutely free of any possible errors whatsoever.”

Comments and syntax texting

The astute reader may have noticed that we have not yet mentioned anything about comments, which may be surprising because comments are typically covered early in the descriptions of most programming languages. WysiScript has an extraordinarily natural and flexible commenting mechanism, but it can seem somewhat puzzling until the rest of the language is understood. Since the artificial link between text and meaning has been discarded, the text of a WysiScript program is free to be used for any comments at all!

For example, the text of the sample WysiScript program in Fig. 4 can be rewritten slightly to help illustrate its structure, as shown in Fig. 6 below.

```
gcd := {  
  (if (a := $1) < (b := $2))  
    then (temp := a, a := b, b := temp)  
    else 0),  
  (if (r := a % b)  
    then eval (a := b, b := r) until (!(r := a % b))  
    else b)  
}  
  
print(gcd(45, 105))
```

Fig. 6. A sample WysiScript program with comments in the text.

In fact, an intelligent WysiScript editor can automate the writing of comments in this way, modifying the text of the program on the fly in order to highlight its syntax. We call this technique “syntax texting,” and as far as we know WysiScript is the first programming language to fully support such a system. Of course, a well-written WysiScript program is usually self-explanatory, but occasionally these textual comments can provide an additional boost to readability.

We understand that this new idea of syntax texting may take some time for traditional programmers to get used to, but we are confident that its benefits will soon become evident. With the availability of syntax texting, we hope that the text of a program will come to be seen as a useful aid in understanding the program's structure.

Conclusions and future work

In this paper we introduced WysiScript, a formatting-based programming language that streamlines the programming process by removing the awkward necessity of writing text. We gave an overview of its main ideas and a tour of its built-in features, demonstrated its simplicity and ease of use with several examples, and proved rigorously that it is equally as powerful as other languages and significantly more colorful. We also proposed syntax texting, a new technique that describes the structure of a program in its own text, which was not previously possible before the introduction of WysiScript.

WysiScript represents the first example of a new and powerful programming paradigm. We believe it is poised to become an important general-purpose language, and we urge its adoption in introductory computer programming courses. In particular, we believe it will appeal to many types of students who may not have an interest in traditional text-based programming, including painters, graphic designers, and the illiterate.

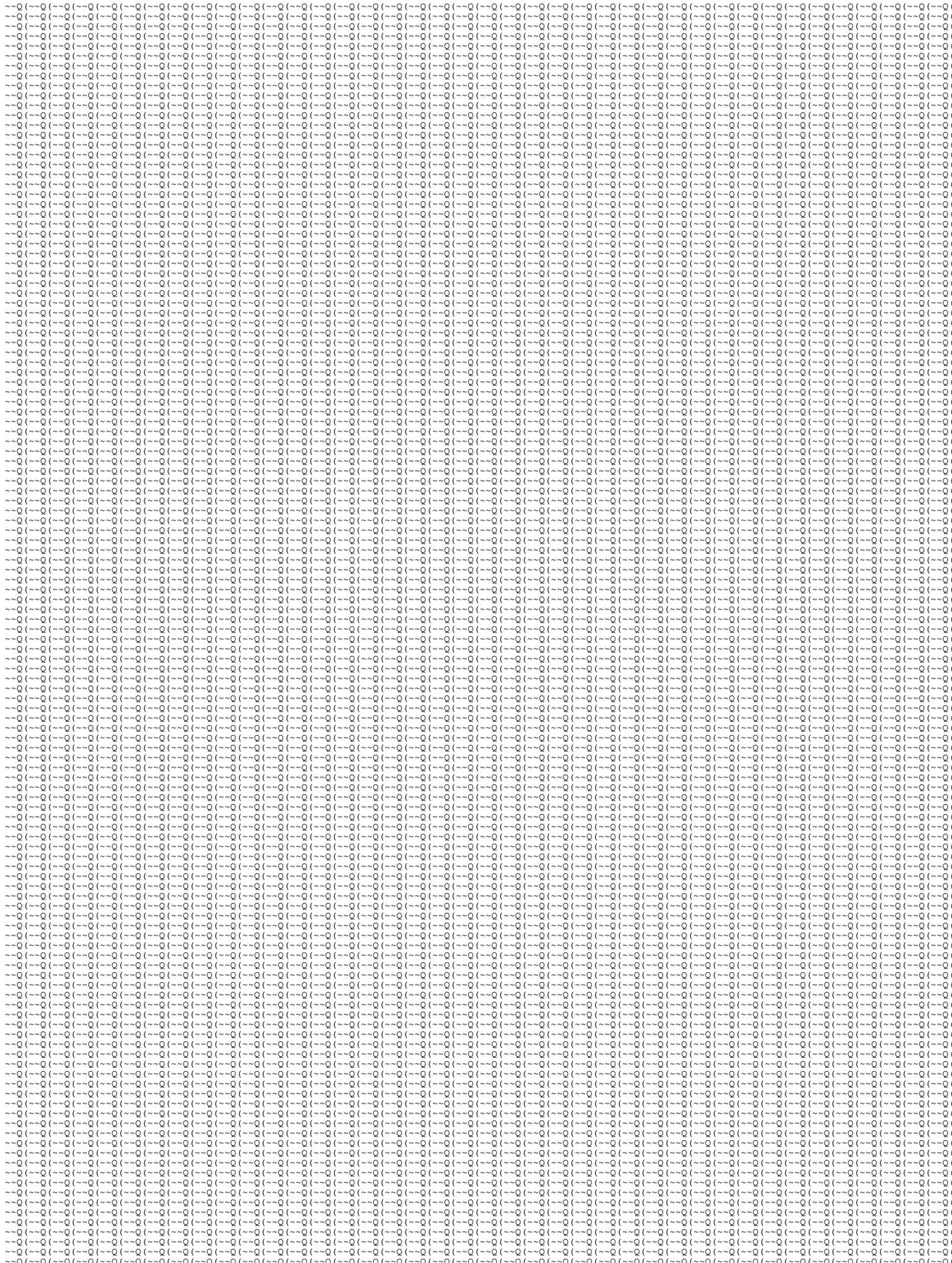
An obvious limitation of WysiScript in its current form is the restriction, imposed by the RGB color model, that a single program can have at most 16,777,216 different user-defined variables and functions. This hinders the use of WysiScript for census-taking in the Netherlands, for example, because the Dutch population is 16,979,729. Future versions of the language may relax this restriction by adding support for additional colors not representable in RGB, such as ultraviolet, infrared, and plaid.

Another promising direction for future research is the extension of WysiScript to object-oriented programming. CSS already provides classes, which can be used as a foundation. With the potential for an object-oriented language in mind, we have reserved the keycolor `thistle` to be used as a reference to the current object.

In order to support quantum computing, which makes use of value superposition and probabilistic algorithms, future versions of WysiScript may make use of the alpha channel in color specifications.

References

- [1] Böhm, Corrado. On a family of Turing machines and the related programming language. *ICC Bulletin*, 3(3):187–194, July 1964.
- [2] Böhm, Corrado, and Jacopini, Giuseppe. Flow diagrams, Turing machines and languages with only two formation rules. *Communications of the ACM*, 9(5):366–371, May 1966.
- [3] Hagenfried, E. W. An analysis of communication in *Megascops roboratus* with applications to computability. *Journal of South American Ornithology*, 26(1):65–74, January 1955.
- [4] Hagenfried, E. W., and Svendt, P. Thermocomputing with diamagnetic flux cores. *Transactions on Circuit Design and Analysis*, 11(4):295–306, October 1958.
- [5] Раджаков, Димитър. Спецификация на компютърни програми чрез навременна натупвам здравата. *Вестник на Българската Академия на Науките*, 43(2):217–242, April 1953.
- [6] Šimunović, Dragana. Grube i glatki predmeti bačeni u kožnoj torbi. *Zbornik Radova Četvrtog Godišnjoj Konferenciji o Računanja s Vrećice i Kutije*, 71–84, December 1957.
- [7] Тураповский, А. М. Новая система для ароматических вычислений с приложениями для производства сыра. *Российский журнал вычислительной химии*, 14(4):320–344, August 1962.



64k of memory, Intel introduced "segments" into the 8086. These are a nightmare for programmers, and when I was a teenager I thought I could perhaps live my whole life without really understanding them. We're back! Roughly speaking, the instruction set allows you to supply 16-bit addresses (offsets), but the processor internally combines these with 16-bit base addresses (segments). The "real address" is (segment * 16 + offset). Some annoying facts:

- The segment registers are changed through different instructions than the regular registers. None are available in printable X86.
- However, we can make some instructions use a different segment register with one of the prefix bytes (e.g. 0x36 makes the next instruction use the SS (stack) segment instead of the default, which might be DS (data)).
- However, some other instructions like PUSH or OUTS can only use a specific segment.
- There are multiple different SEG:OFF pairs that reference the same real address.
- The segment values are not predictable in DOS, because they depend on where DOS happens to place your program.

We'll have to deal with segments for sure, but one consolation (?) is that since we can't change the values, the program will only access the 64k of data within the segments it starts out with.

There are 6 segments, CS (code), DS (data), SS (stack), and three "other" segments ES, FS, and GS.

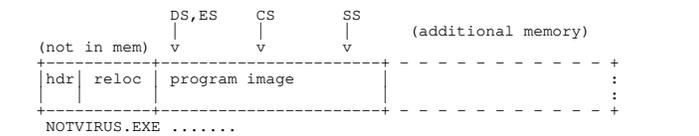
**** 8. Executable file formats, continued... ****

In a DOS .COM file, CS, DS, ES, and SS are all initialized to the same value. This is easy to think about, but it causes a super bad problem for us: The machine stack is inside the same segment as our code. The machine stack is a region of memory that the PUSH and POP instructions use (among others); it starts at the end of this single segment and grows downward (towards lower addresses, where the program's instructions are). If the stack collides with the program, then it will mess up the instructions (which might be an effective way to make self-modifying code, but we don't want to cheat). Most COM programs stay out of the way of the stack by being much smaller than 64k. For good reasons that I will explain later, in this project, execution will need to span the entire code segment. It might be possible to avoid using the stack in our programs, but DOS interrupts (Section 17) are constantly happening as our program runs. These interrupts use the stack, and although they put the stack pointer back where it was and don't modify anything currently on the stack, the values that they PUSH and then POP are still present in memory, overwriting whatever was there. We don't have any way to turn these off, because the CLI instruction ("clear interrupts") is 0xFA, which is not printable. It seems COM files will not work for this project.

This leaves old-style 16-bit DOS EXE files, which do just barely work, and this is what ABC produces. EXE files afford much more flexibility, such as the ability to access up to 640kb (barring tricks) of memory. They also have many features that we do not need or want. An EXE header looks like this:

offset	field	ABC's value (little-endian)	ASCII
00	magic number	0x5A 0x4D	ZM
02	extrabytes	0x7E 0x7E	~~
04	pages in file	0x20 0x23	#
06	relocation entries	0x20 0x20	
08	paragraphs in header	0x20 0x20	
10	minimum memory	0x20 0x20	
12	maximum memory	0x20 0x20	
14	initial stack segment	0x50 0x52	PR
16	initial stack pointer	0x69 0x6E	in
18	checksum	0x74 0x79	ty
20	initial ins pointer	(program dependent)	
22	code segment displacement	0x20 0x20	
24	relocation table start	0x20 0x20	
26	overlay number	0x43 0x20	C

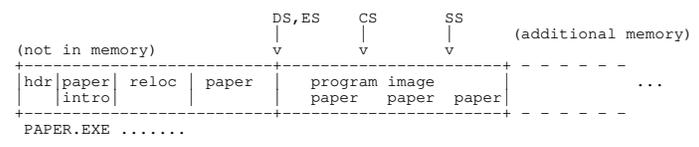
Normally, the header is followed by the relocation table (if any; see below) and then the program image. The program image is some blob of data that gets placed contiguously in memory, with the data segment set to its beginning and the code and stack segments set to wherever the header asks. A typical layout would look like this, with the solid box being the contents of the EXE file:



All of the values in the header are printable, which causes some difficulty. The problem stems from the fact that we must use values that are much larger than is reasonable for several fields; the smallest 16-bit printable number is 0x2020, which is 8224. Several fields are measured in 16-byte "paragraphs" or 512-byte "pages" (anticipating their use in printable executables!), so these values can quickly get out of hand. Naive values cause the program's effective memory requirements to be too large, and DOS does not load our program. Nonetheless, it is possible. The gory details of the solution are documented in exe.sml, but the crux of the solution involves the following tricks:

- Overflow the "pages in file" (a page is 512 bytes, so 0x2320 is 4MB; way beyond the 1MB limit) field to provide a smaller effective value. The file still needs to be pretty big.
- Specify a much larger than usual "pages in header" (0x2020 * 16 = 131kb). Since the header isn't loaded into memory, it doesn't count against the program's memory needs. A really big header also gives us space to store the paper. You're looking at part of the "header" right now.
- Give technically invalid values for some fields (extrabytes, checksum, overlay number); DOS doesn't actually seem to care about these. This helps us get a paper title that's almost readable.

The layout of a compiled program is roughly like this:



PAPER.EXE

This results in an file size of 409,600 bytes, which I believe is the smallest possible. At 160x128 characters per page, this is exactly 20 pages. Since we can't change the segment registers, the active part of our program is only the 64kb data, code and stack segments, and since the stack segment is somewhat unreliable (as described above), we only put stuff in the data and code segments. As a result, we need to be thoughtful about code size; this will be a challenge.

It's not necessary to understand this diagram since you are looking at a 1:1 scale model right now, i.e., the program itself. I'll point these sections out as we encounter them.

**** 9. The Program Segment Prefix ****

The Program Segment Prefix, or PSP, is 256-bytes at the beginning of the data segment. Depending on how you look at it, DOS either overwrites the first 256 bytes of our program image, or the program image is loaded right after it, but starts at address DS:0x0100 rather than DS:0x0000. In any case, we get this for free whether we want it or not, for both COM and EXE files. Since this is just part of DS, programs will be able to read and write the data there. The most useful thing we get is the command line that the program is invoked with from the DOS prompt.

**** 10. Relocations ****

You already saw the header structure (it's the title of the paper) and the relocation table (the full page of "--Q("). For normal programs, the purpose of the relocation table is for DOS to patch the program so that it can know where it's located in memory; each time a program is loaded it might be placed in a different spot. When the program is loaded, DOS goes through all of the entries in the relocation table, and modifies the given location in the program by adding the base segment to the word at that location. Usually this location is part of an instruction sequence like "PUSH imm; POP DS", where imm is some value that we want to be relative to the program's base segment. We can't change segment values, so the relocation table is useless to us. In fact it's harmful, because we have to have 8,224 (0x2020) relocation table entries in order to have a printable header, and whatever offsets are in there will get corrupted when the program is loaded. We repeat the same location over and over, and choose a location that's right after the code segment in memory, a part of the image we don't need. I'll point out the spot that gets overwritten when we get there. The locations are given as segment:offset pairs, which is nice because we have multiple ways to reference a given location. We simply solve for some seg:off such that (seg * 16 + off = addr) and both seg and off are printable.

**** 11. Addressing modes, temporaries, calling convention ****

In any compiler, one must decide on various conventions for how variables are laid out in memory, how registers and temporaries are used, how arguments are passed to functions, and so on. There are lots of such decision in ABC; some are basically normal and some are particular to the weird problems we have to solve. Let's talk about some of the limitations of the instruction set that we have access to, because those inform the low-level design.

In Figure 1, there are several instructions that look like this:

```
AND reg|mod/rm
```

These are each a family of instructions like

```
AND AX <- BX                AND [12345] <- DI
AND BX <- [BP+SI+4]        AND [EBP+12345] <- EBP
```

where the source (on the right) and destination are given by some bits in the instruction's encoding. The instruction always acts between a register and a "mod/rm", with two adjacent opcodes determining whether this is of the form "AND reg <- mod/rm" or "AND mod/rm <- reg". The mod/rm can be one of many possible values; here is a table which you need not absorb:

r16 (/r)	r32 (/r)	AX	CX	DX	BX	SP	BP	SI	DI
		EAX	ECX	EDX	EBX	ESP	EBP	ESI	EDI
Effective Address	Mod R/M	Value of Mod/RM Byte (in Hex)							
[EAX]	00	00 00 08 10	18	*20	*28	*30	*38		
[ECX]	001	01 709 11 19	*21	*29	*31	*39			
[EDX]	010	02 70A 12 1A	*22	*2A	*32	*3A			
[EBX]	011	03 0B 13 1B	*23	*2B	*33	*3B			
[sib]	100	04 0C 14 1C	*24	*2C	*34	*3C			
disp32	101	05 70D 15 1D	*25	*2D	*35	*3D			
[ESI]	110	06 0E 16 1E	*26	*2E	*36	*3E			
[EDI]	111	07 0F 17 1F	*27	*2F	*37	*3F			
[EAX+disp8]	01	000	*40	*48	*50	*58	*60	*68	*70
[ECX+disp8]	001	*41	*49	*51	*59	*61	*69	*71	*79
[EDX+disp8]	010	*42	*4A	*52	*5A	*62	*6A	*72	*7A
[EBX+disp8]	011	*43	*4B	*53	*5B	*63	*6B	*73	*7B
[sib+disp8]	100	*44	*4C	*54	*5C	*64	*6C	*74	*7C
[EBP+disp8]	101	*45	*4D	*55	*5D	*65	*6D	*75	*7D
[ESI+disp8]	110	*46	*4E	*56	*5E	*66	*6E	*76	*7E
[EDI+disp8]	111	*47	*4F	*57	*5F	*67	*6F	*77	*7F
[EAX+disp32]	10	000	80	88	90	98	A0	A8	B0
[ECX+disp32]	001	81	89	91	99	A1	A9	B1	B9
[EDX+disp32]	010	82	8A	92	9A	A2	AA	BA	BB
[EBX+disp32]	011	83	8B	93	9B	A3	AB	BB	BC
[sib+disp32]	100	84	8C	94	9C	AC	BC	CC	CC
[EBP+disp32]	101	85	8D	95	9D	A5	AD	BD	BD
[ESI+disp32]	110	86	8E	96	9E	A6	AE	BE	BE
[EDI+disp32]	111	87	8F	97	9F	A7	AF	B7	B7
AL/AX/EAX	11	000	C0	C8	D0	D8	E0	F0	F8
CL/CX/ECX	001	C1	C9	D1	D9	E1	E9	F1	F9
DL/DX/EDX	010	C2	CA	D2	DA	E2	EA	F2	FA
BL/BX/EBX	011	C3	CB	D3	DB	E3	EB	F3	FB
AH/SP/ESP	100	C4	CC	D4	DC	E4	EC	F4	FC
CH/BP/EBP	101	C5	CD	D5	DD	E5	ED	F5	FD
DH/SI/ESI	110	C6	CE	D6	DE	E6	EE	F6	FE
BH/DI/EDI	111	C7	CF	D7	DF	E7	EF	F7	FF

Figure 2. Addressing modes

The "scaled index byte" (sib) has another table with 224 entries, which we won't get into. There is also a similar, but crazier, table for 16-bit addresses and 8-bit operands. Note that only part of this table is printable (marked with *), which means we can only use a subset of addressing modes. Notably:

- We can't do any register-to-register operations, like "AND AX <- BX". Most compilers use these instructions frequently!
- As a result, exactly one of the source or destination operand is some location in memory.
- The simple addressing modes can only be paired with some registers. For example, AND DI <- [EDX] is allowed, but AND AX <- [EDX] is not. [ESI] means the memory in the location pointed to by the value in the ESI register.

This is even more annoying than x86 usually is. That said, the fact that we don't have register-to-register operations means that register allocation is far less important than usual. Instead, we operate on a set of temporaries, accessed using the [EBP]+disp8 addressing mode. EBP's default segment is SS, so these temporaries are stored in the same segment as the stack. In fact, since we initialized the stack pointer towards the middle of SS (it has to be printable; the maximum value would be 0x7e7e, but we use 0x6e69 to make the title more readable), we have the entire region from that to 0xFFFF to use for temporaries. Each function frame (see below) has its own set of temporaries.

To perform a basic subtraction operation, whereas a traditional compiler is likely to emit an instruction like

```
0x29 0xC2 SUB AX <- DX ; AX = AX - DX
```

ABC emits a sequence like

```
?? MOV AX <- [EBP+0x22] ; AX = tmp2
0x67 0x29 0x45 0x20 SUB [EBP+0x20] <- AX ; tmp0 = tmp0 - AX
```

which is not so bad. (Note that we do not have a MOV instruction; this puzzle is solved below). We often need to do much more work than this to perform a basic operation, and optimization is meaningful (especially things that reduce code size).

The [EBP]+disp8 addressing mode denotes the location in memory at the address in EBP, plus the given 8-bit value (above, 0x22). Note that to encode this mod/rm, we need to write the displacement byte in the opcode, so it must be printable. The EBP register will therefore actually always point 32 bytes before the first temporary, so that temporary 0 is accessed as [EBP+0x20].

With this idea in mind, here is a summary of ABC's low-level design:

- A C pointer is represented as a 16-bit address into the data segment.
- Anything addressable therefore needs to be stored in DS. This includes global variables, local variables and function arguments.
- Global variables are just allocated at compile time to some locations near the beginning of DS.
- A traditional C compiler uses the machine stack to store local variables, but since these need to be in DS, not SS, we maintain a separate stack of arguments and locals in DS, which starts after the global variables and grows towards larger addresses. This is called the locals stack. The register EBX points 32 bytes before the locals stack, so that we can use [EBX+disp8] to efficiently access locals.
- EBP always points 32 bytes before the "temp stack".
- Both stacks (and the machine stack) advance when we make a function call, so that the values of locals and temporaries persist across the function call. ABC only stores the return address on the machine stack.
- Aside from EBX, EBP, and ESP (the machine stack pointer), all other registers can be used for any purpose.

Next, we need to implement a number of low-level primitives that let our program do computation. Let's warm up with something very basic.

**** 12. Putting a value in a register ****

When programming X86 like a normal person, a very common task is to put an arbitrary number (for example, the address of a global, or a value that appears in the user's program) into a register, like

```
0xB8 0x34 0x12 MOV AX <- 0x1234
```

We don't have this instruction available, since its opcode 0xB8 is not printable. Moreover, we need to be able to load arbitrary values, not just printable ones (but the value is part of the instruction encoding).

We do have some ability to load values. For example, we can encode

```
AND AX <- 0x2020
```

since 0x2020 is printable. This clears most of the bits in AX, and then

```
AND AX <- 0x4040
```

will always clear the remainder, since (0x40 & 0x20 = 0x00). With AX containing 0x0000, we could then repeat "INC AX" 1,234 times to reach the desired value. This totally sucks, but it works.

There are often more direct routes. We can XOR and SUB and AND with printable 8- or 16-bit immediate values in addition to INC and DEC. There is probably no "closed form" solution for the quickest route to a given value (the presence of both XOR and SUB makes this rather like a cryptographic function), but we can use computers to help.

We build a routine that generates a series of x86 instructions that load a 16-bit value into AX. In the general case, we do this by loading two 8-bit values and jamming them into AX using a gross trick. To load an arbitrary value into AL (the low byte of AX), ABC uses a table that it creates upon startup. This table is of size 256x256, and gives us the shortest (known) sequence for putting some desired byte DST in AL when AL is known to already contain some byte SRC. This table is populated via something like Dijkstra's "shortest path" algorithm. For starters, the diagonal (SRC = DST) can be initialized to the empty instruction list. We can then use INC and DEC to fill the rest of the table with very inefficient but correct sequences (still, when SRC is 5 and DST is 6, INC AX will remain the best approach!). Next, we maintain a queue of

cells that should be searched (everything goes on the queue except the diagonal, which is already optimal). We repeatedly remove items from the queue and then explore what cells we can reach from that source byte. For example, if we pull out the cell (SRC=0x80, DST=0x01), we try applying XOR, SUB, and AND (with printable immediate values), etc. to the source value 0x80 to see what we get. One such result is that we can get AL=0x00 by doing AND AL <- 0x40. Consulting the cell for (SRC=0x00, DST=0x01), we see that it contains a sequence of length 1 (INC AX), so this gives us a new best solution by concatenating these two paths (AND AL <- 0x40, INC AX), which is much better than (DEC AX, DEC AX, ... 79 times). We iterate this procedure until paths stop improving.

This works well, with only an average of 2.54 bytes of instructions needed to transform a source byte into a destination one (across all possible src/dst pairs). No sequence is longer than 4 bytes. Since this table is big and programmatically computed when the compiler starts, I took some trouble to optimize it (the naive implementation took 13 seconds, which is a bit of an annoying wait every time you run the compiler!). There were a few tricks, but the most fruitful one was to functorize the code that encodes x86 instructions. This code normally works with vectors, and then the test above for the shortest instruction sequence would use Word8Vector.size to compute the best one. In the functorized version, the type of vector is an abstract argument. We instantiate a size-only version of encoding where the "vector of bytes" is actually just the count of bytes, and concatenation is just +. The MLton compiler is then excellent at optimizing this code to throw away the computations of the byte values (they are dead), and this code becomes plenty fast (~800 ms).

The table of instructions contains interesting structure, or at least pretty structure. Since it is 256x256, it can't fit in this paper 1:1, but I cropped to the prettiest part, the leftmost 160 columns. It appears as two full pages in the data segment (Pages 8 and 9) as some cool ASCII triangles. In this graphic, a space character means 0 instructions (this is only the diagonal of course, mainly visible on the first page); '.' means one instruction byte (just INC and DEC, near the diagonal); '-' is two instruction bytes (like XOR AL <- 0x2A); '%' is three; and '#' is four. This fractal pattern (like the Sierpinski triangle?) shows up all over the place in mathematics and computer science and Hyrule. For example it is reminiscent of the matrix of game configurations in k/n Power Hours [KNPH'14].

Once we can load an arbitrary byte into AL, we can fill all of AX with this trick. Suppose that our goal is to load AH=0x12 and AL=0x34. If we don't know anything about AX, we can zero it with two AND instructions. Then we can emit the instructions to load 0x12 starting from the known value 0x00. Then this sequence:

```
instruction AH AL stack (ww, xx, yy, zz stand for
PUSH AX ww 0x12 xx yy zz ... some arbitrary junk)
PUSH 0x3040 ww 0x12 0x12 ww xx yy zz ...
INC SP ww 0x12 0x40 0x30 0x12 ww xx yy zz ...
POP AX ww 0x12 0x30 0x12 ww xx yy zz ...
INC SP 0x12 0x30 ww xx yy zz ...
INC SP 0x12 0x30 xx yy zz ...
```

Remember that x86 is little endian, so the low byte goes on the top of the stack. This trick places two words adjacent on the stack, but then misaligns the stack by doing a manual INC SP (and again at the end to clean up). The result is that AL gets moved into AH, and a known printable value of our choice (0x30 above) into AL. We can then use our table to transform that known value to any desired value into AL, completing the 16-bit value. This is reasonably brief and only touches the AX register, and we use it all the time in the generated code.

**** 13. Moving between registers and memory ****

Another useful kind of instruction is MOV AX <- [EBP+0x20], which moves the 16-bit word at the address in EBP (offset by 0x20) into AX. This is how we read and write temporaries; the "AX <- [EBP+0x20]" part is printable, but we don't have the MOV opcode available (0x89). Fortunately, the XOR instruction is "information-preserving," so it can be used like a MOV. Specifically, if we already have zero in the destination, then XOR *is* a MOV. In order to load from memory we use an instruction sequence like:

```
... various ... set ax <- 0x0000 ; using tricks above
0x67 0x33 0x45 0x20 XOR AX <- [EBP+0x20]
```

To write to memory, we do:

```
0x50 PUSH AX ; save value to write
... various ... set ax <- 0x0000 ; using tricks above
0x67 0x21 0x45 0x20 AND [EBP+0x20] <- AX ; clears to zero
0x58 POP AX ; restore value
0x67 0x31 0x45 0x20 XOR [EBP+0x20] <- AX ; write it
```

This is almost... nice! But don't worry, it gets grosser.

**** 14. Bitwise OR ****

We don't have the OR instruction, but it can be computed with this trick.

1	1	0	0	A
1	0	1	0	B
1	0	0	0	A AND B
0	1	1	0	A XOR B
0	0	0	0	(A AND B) AND (A XOR B)
1	1	1	0	(A AND B) OR (A XOR B)
1	1	1	0	(A AND B) + (A XOR B)
1	1	1	0	A OR B

This is the table of all possible bit combinations that A and B could have; the OR operation is of course only dependent on the pair of bits at each position. First, observe (in your mind; it's not in the table) that A OR B is the same as A + B unless both bits are 1; only in that case do we need to do a carry. So we compute A AND B, and A XOR B; the OR of these two is the same as A OR B (it separates A OR B into the cases where both bits in the input were 1, and the case where exactly one was 1). Since the two expressions never have a 1 bit in the same position, we can compute their OR with +, giving us the desired result. Implementing plus is also a multi-step process, described next:

**** 15. Keeping track of what's up with the accumulator ****

The ABC backend (tactics.sml) generates X86 for some low-level primitives that operate on temporaries, like "Add tmp1 <- tmp2". (This is described in Section 21 when discussing the phases of the compiler.) Because it's expensive to load constants into registers, we go through some trouble to keep track of the machine state as we generate code. This allows us to make some opportunistic improvements. For example, the actual SML code implementing Add on 16-bit numbers looks like this:

```
fun add_tmp16 acc dst_tmp src_tmp : acc =
  let
    val acc = acc ++ AX
  in
    imm_ax16 acc (Word16.fromInt 0xFFFF) //
    XOR (S16, A <- EBP_TEMPORARY src_tmp) ??
    forget_reg16 M.EAX //
    INC AX ??
    forget_reg16 M.EAX //
    SUB (S16, EBP_TEMPORARY dst_tmp <- A) -- AX
  end
```

The approach is to XOR the source value with 0xFFFF and then increment it by 1; this negates the value in two's complement. We can then use the SUB operator, whose opcode is printable, to subtract that negated value, which is the same as adding it. The "accumulator" (variable acc) lets us manage the steps. Without getting into tedious details, "acc ++ AX" claims the register AX so that tactics know not to clobber it; we later return it with "-- AX". The imm_ax16 function loads the value 0xFFFF into AX; this tactic gets to inspect what's known about the machine state. For example, if we happen to have just assembled something that left AX containing 0x0000 (very common) then we can simply DEC AX to get 0xFFFF in one byte. imm_ax16 updates the accumulator to record that AX now contains 0xFFFF, as well as emitting whatever instructions it needs. The // combinator emits a raw instruction, and the ?? combinator allows us to learn or forget a fact about a register. Because some tricks require knowledge of e.g. AL but not AH, the accumulator actually keeps track of each byte of each register independently. It also understands that if you claim ESI, then SI cannot be used (SI is part of ESI), and so on. This is nice, and the semi-monadic syntax allows what looks like assembly code in ML. (Also note the questionable <- and <- (hyphen vs. tilde) datatype constructors that distinguish the two directions of instruction, "reg <- mod/rm" vs. "mod/rm <- reg".) The biggest risk of this approach is if you don't accurately record the state of registers (e.g. you forget to "forget_reg16" after modifying it), because this can lead to tactics making wrong assumptions but only in certain unlucky situations. Some of my worst bugs were from this; it would be cleaner if the accumulator actually simulated the instructions to update its own internal facts, rather than have the programmer make assertions.

Since the accumulator is purely functional, another cool thing we can do is try out multiple different strategies for assembling some block, and pick the best one. For example, when we decrease EBP right before returning from a function (to restore the caller's temporaries), we can either subtract a constant (number of bytes depends on the machine state) or DEC BP over and over (frequently faster).

**** 16. Pointer loads and stores ****

Another primitive we must implement is "load16 dst_tmp <- addr_tmp"; the temporary addr_tmp contains a 16-bit address, and we load the value contained at that address (in DS) and store it in dst_tmp. This is used for pointer dereferencing in the source C program, for example.

It's basically the same as loading from a temporary; we just need to do something like

```
set DI <- 0           ;; macro
XOR DI <- [EBP+0x20]  ;; appropriate addr temporary offset
set SI <- 0           ;; macro
XOR SI <- [DI]        ;; read from the address into SI
set [EBP+0x24] <- 0  ;; appropriate dst temporary offset
XOR [EBP+0x24] <- SI  ;; store it
```

(Again, the syntax [DI] means use the contents of the DI register as a memory address, and load from there. DI's default segment is DS, which is where C pointers always point.) The only complication is that the pure-indirect mod/rm bytes like [DI] can only be paired with certain registers or else they are not printable (Figure 2).

The reason to bring this primitive up is that there's a delightful hack that's possible if the destination temporary and address temporary are the same slot. This situation rarely occurs naturalistically, since it would correspond to unusual C code like (int*)x = (int*)x. However, it is very commonly the output of temporary coalescing (Section 22), since it is typical for the final use of an address to be a load from it. So, this is actually useful (saves about 5% code size), but the main reason to do it is awesomeness! Let's say the single temporary is at EBP+0x20.

```
set DI <- 0           ;; macro
XOR DI <- [EBP+0x20]  ;; load the address into DI.
XOR DI <- [DI]        ;; DI = DI ^ *DI (!?)
XOR [EBP+0x20] <- DI  ;; tmp = address ^ address ^ value
```

The first two steps are reasonable, and put the address into DI. We want to end up with the value (whatever address points to) in the single temporary. Next we execute a crazy instruction, which XORs the address stored in DI with the value it points to. After this, DI contains addr ^ value, sort of like an encrypted version of the value. However, the temporary still contains the address (the "decryption key"), so if we XOR DI into it, we get address ^ address ^ value, which is 0 ^ value, which is just value! It's really nice how short the instruction sequence is, and it only uses a single register. The instruction XOR DI <- [DI] is so weird--it probably occurs in almost no programs, because it is extremely rare for an absolute address to have any relationship with the value it points to. So we get extra style points for finding a legitimate use for it.

Stores are the same idea. The trick actually applies there too, but isn't useful because it doesn't save us instructions, and because it is uncommon for the address and value to be the same temporary in a store operation (store is not really the opposite of load in this sense; both temporaries are read and neither is modified in "Store16 addr_tmp <- src_tmp").

**** 17. Exiting and initializing the program ****

We also want to be able to exit the program when we're done. This is normally done by making a "system call" to an operating system routine

to tell it that we're done and the program can be unloaded. In DOS, you make system calls by triggering a processor interrupt with the INT instruction, which is a way of telling the operating system, "Check this out!" We don't have access to this instruction, whose opcode is 0xCD. Alas! The INT instruction is a gateway to all sorts of useful functionality, like printing strings and reading from the keyboard, reading and writing files, changing video modes, and so on, so it's very sad to go without it. (The EICAR test virus uses self-modifying code to create two INT instructions; one is to print the string and the second is to exit.) In DOS, INT 0x21 is the most useful one; you set registers to some values to access dozens of different functions.

INT 0x21 is so common that it appears in the Program Segment Prefix that's always loaded at the beginning of the data segment. It's just sitting there amidst some zeroes:

```
...
DS:0x004A  0x00 0x00  ADD [BX+SI] <- AL
DS:0x004C  0x00 0x00  ADD [BX+SI] <- AL
DS:0x004E  0x00 0x00  ADD [BX+SI] <- AL
DS:0x0050  0xCD 0x21  INT 0x21
DS:0x0051  0xCB          RETF
DS:0x004E  0x00 0x00  ADD [BX+SI] <- AL
DS:0x004E  0x00 0x00  ADD [BX+SI] <- AL
...
```

It even tantalizingly has RETF (far return from function call) immediately after it, like it was planned there by some puzzlemaker of years past, exactly for this kind of situation. (I don't actually know why it's there!) RETF pops both a return address and return segment, so if we could manage to put a return address on the stack (not hard) and the code segment (we don't know it, but we could probably use the relocation table to write it somewhere) beneath it, and then somehow transfer control to DS:0x0050, we'd have a fully general INT 0x21 to use! It would even help with the loop problem (next section) since it lets us return to an arbitrary address, and could conceivably even let us escape the confines of always executing code within the initial code segment CS (because RETF modifies CS). But speaking of confines, none of this will work, because we have no way of modifying CS to start executing code out of DS. Too bad, so sad. (This idea might pan out for a COM file where CS=DS, but there we have no relocation table so figuring out what segment value to put in the stack would require some other hack. We also have the Loop problem, preventing us from reliably jumping to DS:0x0050. Might be worth further exploration.)

Jumping the program to a non-printable instruction is also a bit questionable, though it's not an instruction that we wrote there, so this does not violate our self-modifying code fatwa. Is it wrong for a waiter to serve the ovo lacto vegetarian with vegetarian food that causes him to eat non-vegetarian food that the customer himself brought with him? Who can say?

This is not hopeless. The way interrupts actually work is to stop the current execution (saving the state of the registers on the stack) and then consult a table of "interrupt vectors" (in my opinion the table itself should be called the "interrupt vector", containing addresses) at the address 0x0000:0x0000 (i.e., right at the beginning of memory). Each interrupt has a number, and each address is a 32-bit segment:offset pair. So the address at 4 * 0x21 = 0x0084 is the location of DOS's code for INT 0x21. In 16-bit real mode programs, there's nothing special about the operating system; you can just jump directly into it if you want, or overwrite it with your own stuff. In fact, this is how many viruses work; for example by replacing the address for INT 0x21 with their own code, and intercepting file operations to insert viruses before calling through to the original INT 0x21 handler so that everything still works.

Fetching the INT 0x21 address is not immediately useful, because we can't transfer control to it; we don't have the CALL instruction. In fact, the only JMP instructions we have must jump a small fixed distance forward (next section). But! The INT instruction is not the only way to trigger interrupts. The timer interrupt is firing continuously, messing with our stack, for example. We can modify the interrupt vector table to make the timer interrupt (INT 0x8) instead point to the INT 0x21 code, and then "wait" for a timer interrupt to happen, and maybe restore the old timer interrupt code when we're done. This might work, but it seems extremely brittle. (Also, the timer interrupt handler has to perform certain low-level duties or else the system will freeze.) Fortunately there's a better choice: The CPU will also trigger an interrupt when an illegal instruction is executed. Normally the illegal instruction handler would do something like crash the program gracefully (in Unix, it sends the SIGILL signal. Sadly there is no SIGBOVIX.) Do we have an illegal instruction inside printable x86? In fact we do!

0x63 Adjust RPL Field of Segment Selector

... it's just sitting in there, this totally weird instruction with no other possible uses amidst a bunch of sensible ones. This instruction is for some operating system privilege stuff, and is illegal in real mode.

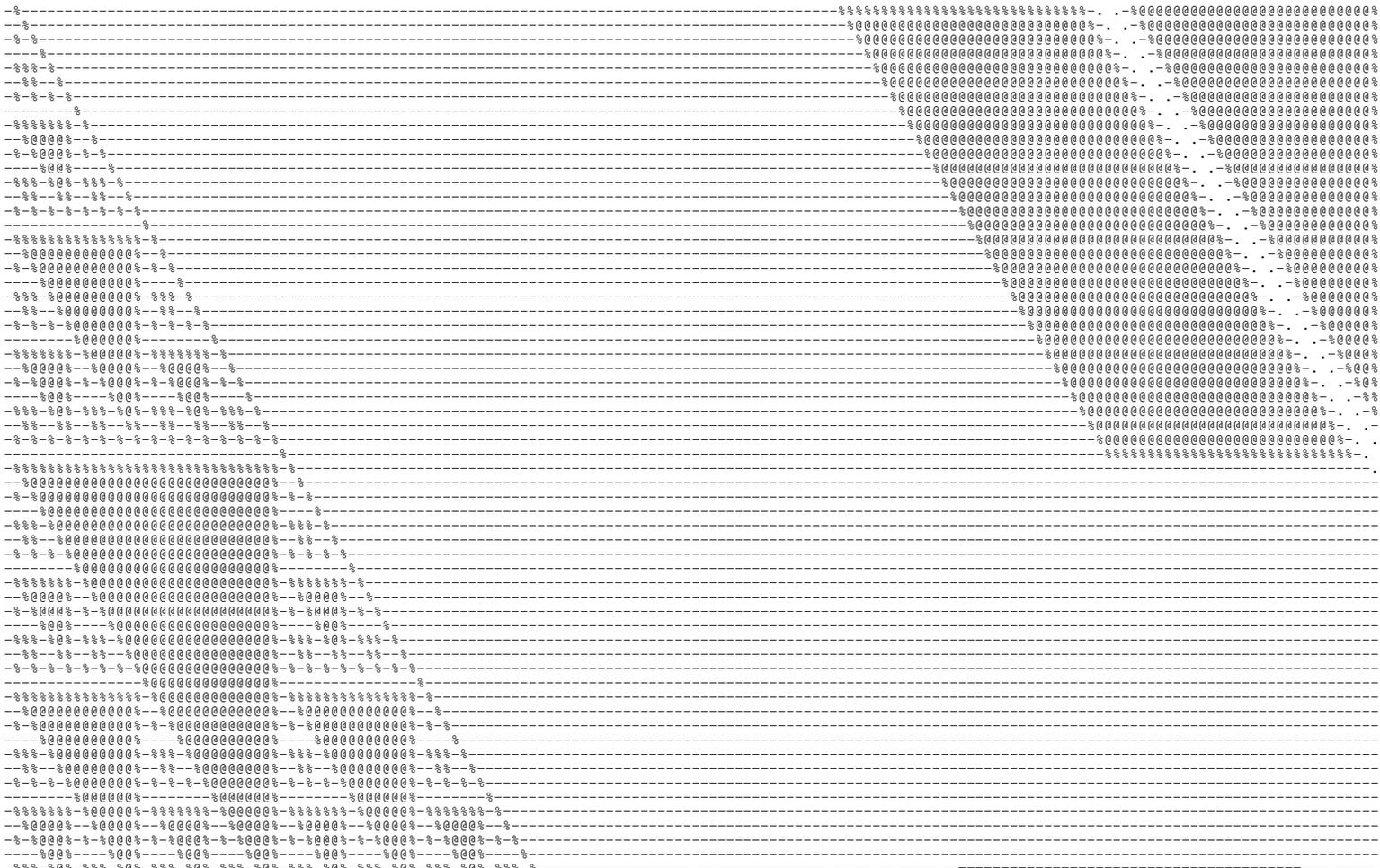
So, when we first start up an ABC program, one of the first things we do is read the address of the INT 0x21 handler at 0x0000:0x0084, and write it over the INT 0x06 (illegal instruction) handler. Luckily the FS segment is set to 0x0000 when our program starts (we can't change it), so we can use the FS segment override instruction to access the beginning of RAM. Once we overwrite the address, then whenever we want we can set up argument registers for the system call "exit" (AH = 0x4c, AL = status code), and execute the illegal ARPL instruction. This will trigger interrupt 0x06, which is now actually the INT 0x21 code, and DOS will "cleanly" exit the program for us.

It is very tempting to use this trick to make other system calls through INT 0x21, or perhaps to jump to arbitrary addresses of our choosing! Sadly, there are two very serious issues:

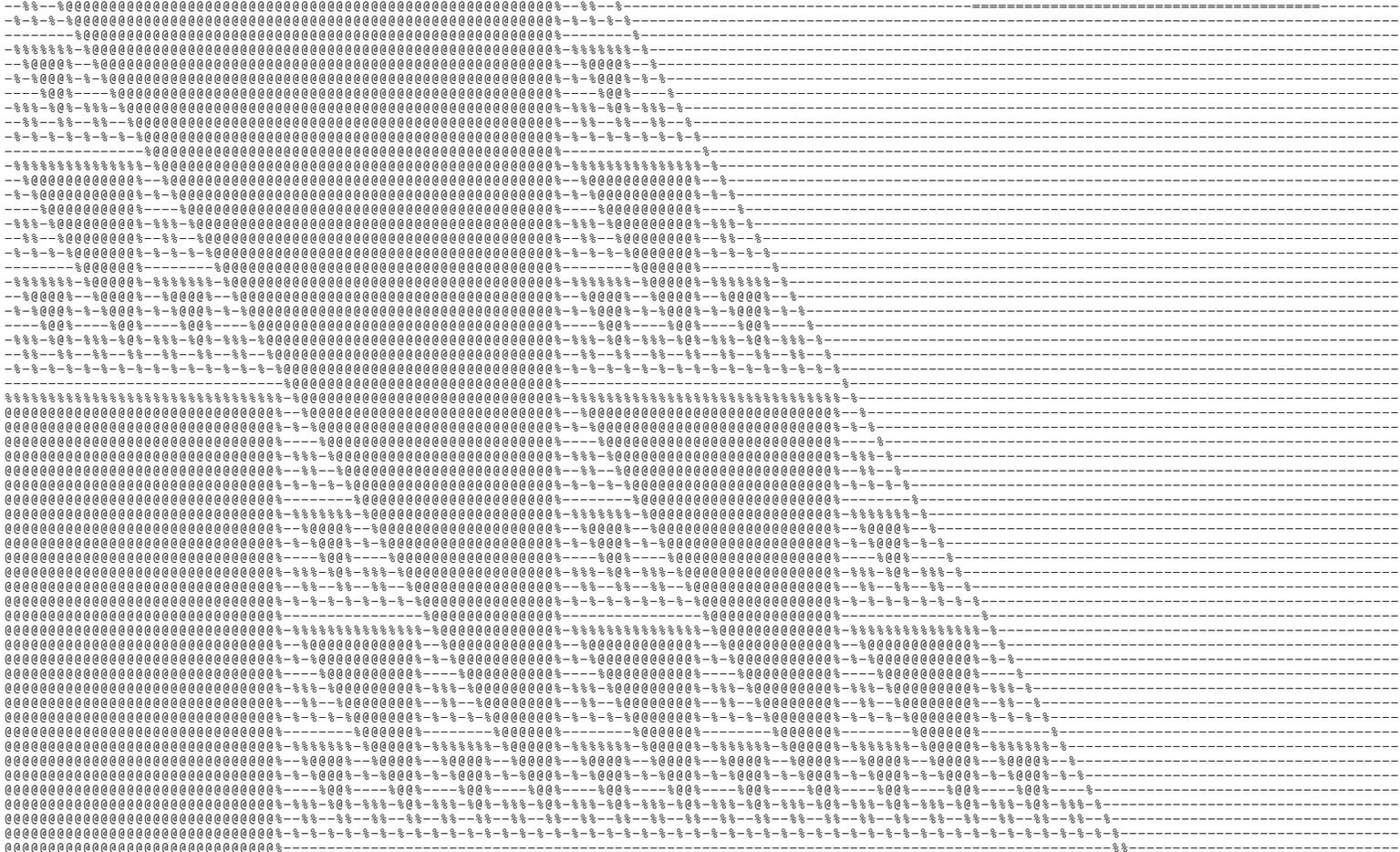
- When the processor triggers the illegal instruction interrupt, the return address that it pushes on the stack is the address of the illegal instruction itself, not the one that follows it. So when the interrupt handler returns, it simply executes another illegal instruction.
- When the interrupt is triggered, it clears the interrupt flag (so that for example the timer interrupt doesn't fire while it's already running). Only a few instructions, which we don't have access to, can restore the interrupt flag. This means that we would only be able to do this once, and after we did, many things would stop working because interrupts would stop firing.

Neither of these issues are a problem for the exit system call, since we only exit once. YOEO!

This is part of the data segment. What else to put in the data segment but some data? You're looking at the data now.



This pretty picture is the number of instruction bytes needed to change a given 8-bit value in the AL register to some other desired value. It's discussed in the section called "Putting a value in a register."



. Anyway, one backwards jump is enough! We can set things up so that
. whenever we need to jump backwards, we instead jump forward until we're
. at the end of the segment, then jump across that boundary (overflowing
. back to the beginning) and then keep jumping forward until we get where
. we need to be. This is delicate, but it works.

. One other issue with jumps is that we can only jump a fixed distance;
. there is no equivalent to "MOV EIP <- AX" to jump to a computed
. location. We need this functionality to implement two C features:
. Function pointers (the destination of a function call is not known at
. compile time) and returning from functions (the function can be called
. from multiple sites, so we need to know which site to return to).

. ** 19. The ladder **

. To solve the various problems with jumps, we build the program around
. what's called a "ladder" in the code. The whole program is broken up
. into small blocks of code. Each one is given a sequential "number" (this
. has nothing to do with the memory location, just its sequence in the
. list of blocks). Each block starts with a "rung," which is the following
. code

```
.     DEC SI  
.     JNZ +disp8
```

. where disp8 is a printable displacement that brings us downward to the
. next block. We decrement the SI register to count down to the block we
. want, and if it is Not Zero yet, then we jump to the next one. If zero,
. we execute the block. Inside a block, if we ever want to perform a jump
. to some arbitrary block dest_block, then we can compute:

```
.     offset = (dest_block - current_block) mod num_blocks  
.     si = (if offset = 0 then num_blocks else offset)  
.     jmp to next rung
```

. Every block knows its current number, so the offset is just a constant.
. Note that the destination block's number may be before the current
. block, which is why we need to mod by the total number of blocks
. (yielding a non-negative result). SI cannot be zero, because the first
. thing we do is DEC it, so a self-loop requires setting to num_blocks, a
. full cycle.

. To perform a jump to a code location not known at compile time (e.g.
. from a return address (block number) on the stack, we can just perform
. the same computation as above. We do not have an efficient mod operation

(implementing it seems to need loops, in fact, a circularity!), so
. instead we actually compute (dest_block - current_block) + num_blocks.
. This is always positive as needed, but requires forward jumps to make an
. entire cycle around the entire ladder ("Turn the dial to the left,
. passing zero and the first number...").

The blocks are laid out sequentially in the program until we get too
. close to the end of the segment; when we do, we make sure to perform an
. unconditional jump across the segment boundary, wrapping around. This
. jump need not DEC SI. In fact, most programs do not fill the entire code
. segment, so we end up padding the end and beginning of the segment with
. jumps to span the unused space. For these padding jumps, we definitely
. don't want to DEC SI, both because that's more instructions to execute,
. and because we don't know the amount of padding ahead of time (see the
. section on Assembling below).

There are many annoyances! A jump cannot be too short (less than 32
. bytes) or too long (127 bytes). The viable range is large enough to
. build nontrivial programs, but is a significant constraint for us.

We don't have access to a non-conditional JMP instruction. There are a
. few tricks for simulating it. When computing a jump to a known label, we
. can just know the state of flags because we've just performed some
. computation. Even when doing a jump to a computed block number, we know
. that the result of subtraction is not zero, so we can always use the JNZ
. instruction. Occasionally we need to do a jump without knowing our state
. at all. XOR always clears the Overflow flag, so something like

```
.     XOR AX <- [DI]  
.     XOR AX <- [DI]  
.     JNO disp
```

keeps AX unperturbed and always performs the jump. A little shorter is

```
.     JNO disp  
.     JO (disp - 2)
```

which jumps to the same target whether the Overflow flag is set or not,
. but is more annoying because we need to keep track of two displacements.

. ** 20. Assembling **

Assembling the program is the process of generating actual instruction
. bytes (here, printable x86) from some semi-abstract representation of
. instructions (in ABC, this is the LLVMNOP language discussed in the next
. section). Assembling has a self-dependency: In order to generate

instructions like jumps and loads of addresses, the assembler needs to know where code is located. But in order to know where code is located, the assembler needs to generate it. In most assembler tasks, this is reasonably straightforward: When we need to generate an instruction like "MOV AX <- offset data", we just emit "MOV AX <- 0x0000" and save for later an obligation to overwrite the zeroes with the address of "data", once we know where we placed it. This works because the encoding of the MOV instruction is the same length no matter what 16-bit value we load. The same holds for JMP instructions (with the caveat that smart assemblers can JMP+disp8 for nearby labels and JMP+disp16 for further ones; these instructions have different lengths) and others.

For the ABC compiler this step is quite bad:

- Loading any immediate value has a length ranging from 0 bytes (it's already in the register) to like 16. It's dependent on both the value being loaded and the context (contents of registers).
- The rungs that start each code block must be able to Jcc+disp8 all the way to the next block. This jump distance can't be too big, or else it can't be encoded (or is not printable).
- Jumps within a block always target the next block, but the jump distance can't be too short (or the displacement byte is not printable).
- Since blocks are numbered sequentially and relative addresses are computed modulo the total number of blocks, logical code addresses depend on the number of blocks and their order.

As a result, assembling is an iterative process. We take the program's blocks and translate them into position-independent machine code. One positive thing about the printable, non-self-modifying subset of x86 is that none of the instructions actually depend on what address they're placed at (except perhaps a Jcc instruction used to overflow the instruction pointer). Still, we don't know even the relative location of the next block yet, so we also record the offset of the displacement byte for any Jcc instruction we emit.

Next, we take these blocks and attempt to allocate them into the code segment. This can fail for the reasons above, usually after we've placed a block far enough from the preceding one that all jumps in the first are printable (at least 0x20 bytes), the rung at the beginning of that block can't target the second (because it is more than 0x7e+0x03 bytes away). We gather all such problem blocks and bisect the LLVMNOP code into two smaller blocks. Then we try again. When we succeed, we can fill in the displacement bytes for the Jcc instructions to create valid printable code. There are various opportunities to be smarter about this (for example, bisecting the LLVMNOP assumes that all such instructions assemble to the same length, which is not remotely true); tox86.sml contains several ideas.

Since the initial instruction pointer must be printable, we start laying out blocks towards the middle of the code segment. If a block would run off the end of CS, then we need to pad that region with jumps that get up close to the end of the segment and then do an overflowing jump past CS:0xFFFF before continuing layout. Once we run out of blocks, we also need to pad any remaining code space with jumps in order to bring control back to the first rung, since the ladder needs to be a complete cycle in order to work. It's easy to pick out the texture of this padding in the code segment (e.g. pages 14, 16).

** 21. LLVMNOP **

Knowing our low-level endpoint, I can now work backwards through the compiler. The compiler generally proceeds by a series of intermediate languages, the last of which is called LLVMNOP.

This language is an assembly-like language that has explicit *data* layout, but not not explicit *code* layout. By that, I mean that every function knows the size and offset of its locals and arguments in the current local frame, and the size and address of each global variable is known, as well as the global's initial values (if printable). It is akin to LLVM [LLVM'04], but doesn't really have anything to do with it. LLVM is an excellent tool for writing compilers (superficially, it looks like a good way to write a new C compiler targeting an architecture like printable x86!) but isn't really suitable for this project because it assumes that the output architecture has certain standard operations efficiently available, which is frequently not the case for printable x86.

A sample of LLVMNOP constructs are:

```
cmd ::= Add tmp <- tmp
      Xor tmp <- tmp
      Push tmp
      Pop tmp
      Mov tmp <- tmp
      Immediate16 tmp <- word16
      Load16 tmp <- tmp
      Store16 tmp <- tmp
      Load8 tmp <- tmp
      Store8 tmp <- tmp
      ExpandFrame i
      PopJumpInd
      JumpCond cond, label
      ...
      Out8
      Init
      Exit

cond ::= Below tmp, tmp
       BelowEq tmp, tmp
       ...
       EqZero tmp
       True
```

LLVMNOP exists in both a "named" and "explicit" version. In the named version, temporaries (tmp) are strings paired with a size (16 or 32 bits). In the explicit version, temporaries are given as a size and offset from the current temporary frame (EBP). The named version is transformed to the explicit version by the process called Allocation (below).

Commands are basically assembly instructions that we might have in a more expressive architecture; note for example that we have Add, which is not native in printable x86 (we implement it by computing the two's complement negation, and then subtracting). Even commands that have a corresponding printable x86 instruction like XOR are still compiled into multiple opcodes, since they read and write arguments to temporaries, not registers. We discussed the implementation of operations like Load16, Immediate16, and Mov in a previous section.

A program consists of a series of labeled blocks. JumpCond pairs a condition (signed and unsigned comparisons, etc.) with a jump to a label. The possible conditions map to the Jcc instructions that we have available. Since opcode 0x7F (Jump Greater) is not actually printable, all of the conditions "face less;" the condition Greater(A, B) is equivalent to Less(B, A). An earlier phase does this rewrite. Also note that in C, a < b is an expression that can be used in any context, not just for control flow; here the comparison is inextricably linked to a jump, since CMP only sets FLAGS, and FLAGS can only be used for jumping. An earlier phase removes the expression forms as well, without being too wasteful when the programmer writes "if (x < 1)" to begin with.

The only way to jump to a non-constant destination is with PopJumpInd, which is basically the RET assembly instruction. It pops an address (block number) from the top of the machine stack, and unconditionally transfers control to that label (by computing the number of blocks to traverse, then jumping to the ladder). This is indeed used to return from a function call, as well as to call a function through a function pointer. It takes its argument on the stack (as opposed to using the existing "Pop tmp" and then "JumpInd tmp") because while we're setting up a function call, we need to move the temporary frame pointer, after which point it is unsafe to access temporaries. The stack, however, is a stable place to stash data.

Since we have some higher-level operations like Mov available, we can implement some delicate maneuvers like function calls as sequences of multiple commands. On the other hand, for some primitives like Init and Exit, there's no real value in breaking them into smaller pieces. Some other complex primitives like Out8 have no analogous feature in C; these are provided as sort of "intrinsic" that can be used to do low-level programming in C. We'll discuss Out8 in Section 27 when we talk about IO. Other primitives, such as one called "Argv" that is used to initialize the argv parameter to main during initialization, is compiled away when we convert to LLVMNOP. In this case, the Argv primitive just creates a global array containing two elements: The second is zero ("null") as required by the standard, and the first is the constant address 0x0081, which is a pointer into the Program Segment Prefix where DOS stores the command line (untokenized; the programmer must do any processing she desires).

** 22. Temporary allocation **

Temporary allocation is fairly standard. We use a dataflow-based liveness calculation to determine which temporaries interfere with one another; if two temporaries of the same size don't interfere, then they can use the same slot, so they are coalesced into one. We prioritize coalescing temporaries in a "Mov tmp1 <- tmp2" so that we get the no-op instruction "Mov tmp1 <- tmp1"; this is possible for a great many Movs, and allows us to be much more regular in the phase that generates LLVMNOP without compromising code size. We then prioritize temporaries that appear in a "Load16 tmp1 <- tmp2" instruction since we have a nice trick for that one when both are the same. After that, we just greedily coalesce temporaries until it is no longer possible. Fancier register allocation techniques like graph coloring would work here (this part of the compiler is very traditional), but there's not much need: We have over 40 16-bit temporaries, all of which are just as efficient to access, so we mainly just want to keep the total number used small so that EBP offsets are printable. Having a smaller temporary frame size allows deeper recursion, as well.

The compilation strategy ends up storing almost all immediate results in temporaries, which is not that suboptimal since all operations need to be between a register and memory anyway. However, many pairs of instructions could keep a just-computed value in a register rather than bothering to write it. This is not yet implemented, but the idea is that we could introduce a small number of registers (probably just one?) in addition to the numbered temporaries, and use those in the output of Allocation. This could produce significantly closer to hand-written code, without the need to change much in the backend.

** 23. CIL **

The intermediate language that precedes the named LLVMNOP code is called CIL, for C Intermediate Language. It's intended to be a desugared and more explicit version of C. Some examples of the of CIL grammar:

```
signedness ::= Signed | Unsigned

type ::= Pointer type
       Code type, type list
       Word32
       Word16
       Word8
       ...

builtin ::= B_EXIT | B_ARGC | B_ARGV | B_PUTC | B_OUT8

value ::= Var v
        AddressLiteral loc, type
        FunctionLiteral name, type, type list
        Word8Literal w8
        Word16Literal w16
        Word32Literal w32

exp ::= Value value
       Plus width, value, value
       LessEq width, value, value
       Load width, value
       Promote width, width, signedness, value
       Call value, value list
       Builtin builtin, value list
       ...

stmt ::= Bind v : type = exp in stmt
       Store width value = value in stmt
       GotoIf cond, string, stmt
       Return value
       ...
```

And lots more stuff. A program is a collection of functions, each of which is a collection of named statements (the stmt type is recursive, with a single statement representing a series of C statements until we reach a Return or unconditional Goto). Programs also have a set of globals with initialization code for them. Note that CIL has ML-style lexically scoped variables which are only in scope for the given block. Since C's semantics for variables allow them to be addressed and modified, we convert all C variables into explicit loads from and stores to memory.




```

.....
// ABC provides no standard library, so you gotta roll
// your own.
int strlen(unsigned char *s) {
    int len = 0;
    while ((int)*s != (int)0) {
        len++;
        s = (unsigned char *)((int)s + (int)1);
    }
    return len;
}

// DOS command lines always start with a space, which is annoying.
// Strip that. DOS also terminates the command line with 0x0D, not
// 0x00. This function updates it in place so that we can use normal
// string routines on it.
int MakeArgString(unsigned char **argstring) {
    unsigned char *s = *argstring;
    while (*s == (int)' ') {
        s = (unsigned char *)((int)s + (int)1);
    }
    *argstring = s;

    while ((int)*s != (int)0x0D) {
        s = (unsigned char *)((int)s + (int)1);
    }
    *s = (unsigned char)0;
    return 0;
}

// We pick octave 4 as the base one; this is fairly canonical and
// benefits us since this array is all printable. Note that A4 is
// higher than C4, since octave 4 begins at the note C4. This
// array maps A...G to the corresponding MIDI note.
unsigned char *octave4 =
    "9" // A = 57
    "8" // B = 59
    "0" // C = 48 = 0
    "2" // D = 50
    "4" // E = 52
    "5" // F = 53
    "7" // G = 55

// Parse a character c (must be capital A,B,C,D,E,F,G)
// and interpret any suffixes as well.
int ParseNote(unsigned char *ptr, int c, int *idx) {
    int midi;
    int offset = c - (int)'A';
    int nextc;
    midi = octave4[offset];
    for (;;) {
        nextc = (int)ptr[*idx];
        switch (nextc) {
            case '^':
                // Up octave.
                midi += (int)12;
                break;
            case ',':
                // Down octave.
                midi -= (int)12;
                break;
            default:
                // Not suffix, so we're done (and don't consume
                // the character.)
                return midi;
        }
        *idx = *idx + (int)1;
    }
}

unsigned int ParseLength(unsigned char *ptr, int *idx) {
    int c = (int)ptr[*idx];
    if (c >= (int)'2' && c <= (int)'8') {
        int m = c - (int)'1';
        *idx = *idx + (int)1;
        return (unsigned int)2048 * m;
    }
    return (unsigned int)2048;
}

// Parse the song description (ptr) starting at *idx. Updates *idx to
// point after the parsed note. Updates *len to be the length in some
// unspecified for-loop unit. Returns the MIDI note to play next, or 0
// when the song is done.
int GetMidi(unsigned char *ptr, int *idx, unsigned int *len) {
    int c, midi_note;
    int sharpflat = 0;
    for (;;) {
        c = (int)(ptr[*idx]);

        // End of string literal.
        if (c == (int)0) return 0;
        // End of command-line argument.
        if (c == (int)0x0D) return 0;

        // Advance to next character.
        *idx = *idx + (int)1;

        switch (c) {
            case '^':
                sharpflat++;
                break;
            case ',':
                sharpflat--;
                break;
            case '=':
                // Nothing. We assume key of C, so there are no naturals.
                break;
            case 'z':
                *len = ParseLength(ptr, idx);
                // No sound.
                return 128;
            default:
                if (c >= (int)'A' && c <= (int)'G') {
                    midi_note = ParseNote(ptr, c, idx) + sharpflat;
                    *len = ParseLength(ptr, idx);
                    return midi_note;
                } else if (c >= (int)'a' && c <= (int)'g') {
                    midi_note = ParseNote(ptr, c - (int)32, idx) + (int)12 + sharpflat;
                    *len = ParseLength(ptr, idx);
                    return midi_note;
                }
        }
    }
}

int main(int argc, unsigned char **argv) {
    int song_idx = 0, j, midi_note;
    unsigned char *cmdline = *argv;
    unsigned char *song;
    MakeArgString(&cmdline);

    // First test for known songs. After that, if we have a command line,
    // use it. Otherwise,
    if (streq(cmdline, (unsigned char *)"-alphabet")) {
        song = alphabet;
    } else if (streq(cmdline, (unsigned char *)"-plumber")) {
        song = plumber;
    } else if (streq(cmdline, (unsigned char *)"-bluehair")) {
        song = bluehair;
    } else if (strlen(cmdline) > (int)0) {
        song = cmdline;
    } else {
        song = default_song;
    }

    Quiet();

    // Initialize the Adlib instrument.
    Adlib((int)0x20, (int)0x01); // Modulator multiple 1.
    Adlib((int)0x40, (int)0x10); // Modulator gain ~ 40db.
    Adlib((int)0x60, (int)0xF0); // Modulator attack: quick. Decay: long.
    Adlib((int)0x80, (int)0x77); // Modulator sustain: med. Release: med.
    Adlib((int)0x23, (int)0x01); // Carrier multiple to 1.
    Adlib((int)0x43, (int)0x00); // Carrier at max volume.
    Adlib((int)0x63, (int)0xF0); // Carrier attack: quick. Decay: long.
    Adlib((int)0x83, (int)0x77); // Carrier sustain: med. release: med.

    for (;;) {
        unsigned int len;
        midi_note = GetMidi(song, &song_idx, &len);
        if (midi_note == (int)0) break;
        PlayNote(midi_note);
        for (j = (int)0; j < len; j++) {}
    }

    Quiet();
    return 0;
}

** 30. Is this useful for anything? **

No. This is a SIGBOVIK paper. <3

** 31. Future work **

There are many code size optimizations possible, and while nontrivial
programs can fit in 64k (such as the one in this paper), larger ones
will run up against that boundary quickly. Probably a factor of about
4 can be gained through a few hard but straightforward optimizations.
Can we break free of the 64k boundary? Earlier we noted that when
execution exceeds CS:0xFFFF, it simply continues to CS:0x00010000
unless a jump is executed across that boundary; this address is
pointing to bytes that are part of our program image (this text is
there, in fact), so conceivably we could write code here. One
significant issue is that interrupts, which are constantly firing,
push 16-bit versions of CS and IP onto the stack, and then RETF
(return far) to that address. This means that if an interrupt happens
while we are executing in this extended address space, we will return
to CS:(EIP & 0xFFFF). If we had control over interrupts, this might be
a good way to return to the normal 16-bit code segment (i.e., to
perform a backwards jump), but as discussed, we do not. We may be able
to globally suppress interrupts, like by using our single illegal
instruction interrupt during initialization, with the interrupt
handler pointing just to code that we control (and never returning
from it). This leaves the interrupt flag cleared, as discussed. The
computer will be non-functional in many ways, because the operating
system will no longer run, but we might still be able to do
rudimentary port-based I/O, or build our own non-interrupt-based OS.
With interrupts suppressed, we can't use the interrupt trick to return
to CS:0000. However, my reading of the Intel manual [INTC] seems to
imply that a jump performed from this region can be forced into 16-bit
mode (thus being subject to the & 0xFFFF overflow) with an address
size prefix; however, this does not seem to be the case in DOSBox.
Given how unusual this situation is, it may even be a bug in DOSBox's
CPU emulator. Having access to a full megabyte of code (it still needs
to fit in the EXE container) would be exciting, since it would allow
us to build much more significant systems (e.g. standard malloc and a
floating point emulator); more investigation is warranted here.

I initially designed CIL with the thought that it could be used for
multiple such "compile C to X" projects. These are primarily jokes,
but can occasionally be of legitimate use for low-level domain-
specific tasks where the existence of a reasonable and familiar
high-level syntax pays for the effort of writing a simple backend.
(When making such a decision I like to also weight the effort by the
enjoyment of each task: i.e., the cost is like

(1 - fun of writing backend) * time writing backend vs
(pain of writing low-level code by hand) *
time writing low-level code by hand

... but I have been informed that not all computer work is done purely
for fun.) This "portable assembler" application of C remains relevant
today, and CIL or LLVMNOP is a much simpler than GCC or LLVM.

Anyway, I discovered that the design of such a thing is not so easy.
While it is possible to "compile away" certain features by turning
them into something "simpler," it's not straightforward what feature
set to target. For example, for ABC, we compile away the | operator
into &, ^, -, and +1. In another setting, | may very well be present
instead of &. We normally think of the >> and << shift operators as
being fundamental, but in ABC they are inaccessible. I find the
expression forms like "a < b" much easier to think about than the
combined test-and-branch version, but the latter is much better when
targeting x86, and important for producing reasonable code in ABC. I
do think it would be possible to develop a simple and general language
for this niche where certain constructs could be compiled away in
favor of others, at the direction of the compiler author, but such a
thing is firmly future work.

On the topic of taking away, one might ask: What is the minimal subset
of bytes we could imagine using?
.....

```

There are some trivial subtractions: We never emit the BOUND instruction (0x62, lowercase b) and it does not seem useful; a few of the segment prefix instructions are also unused. The instructions like "ASCII Adjust After Addition" are currently unused, but since they act on AX in a predictable way, they could provide ways to improve the routines to load immediate values. But we're talking about reducing the surface, not increasing it. And speaking of loading immediate values, we do certainly make use of the entire set of printable bytes in these routines (as arguments to XOR, SUB, PUSH, etc.), but on the other hand, we can also reach any value from a known starting point by INC and DEC, taking at most 0x7FFF instructions (half the size of the code segment, unfortunately). More essential is our ability to set a register to a known value, which today requires two or more printable values whose bitwise AND is 0. Sadly, though we could go through some pains to remove bytes from the gamut here and there, no natural subset like "only lowercase letters" or "alphanumeric" jumps out as a straightforward extension; we rely on the control flow in the late lowercase letters (Jcc) and the basic ops in the early punctuation (AND/XOR), not to mention that the EXE header barely works within the existing constraints with access to both "small" (0x2020) and "large" (0x7e) constants.

Others have produced compilers for high-level languages with very reduced instruction sets. In an extreme case, Dolan shows [MOV'13] that the mov instruction on its own is Turing-complete (note however that this requires a "single absolute jump" to the top of the program, an issue similar to what we encounter in printable x86, only we do not cheat by inserting any out-of-gamut instructions). Another enterprising programmer, Domas, implemented a C compiler that produces only MOV instructions [MVF'16]. I didn't look at it while writing ABC (spoilery!) but he avoids using any JMP instruction the same way that I exit the program (generating illegal instructions but rewriting the interrupt handler). While awesome, the problem is somewhat different from what ABC solves; here we are fundamentally concerned with what bytes appear in the executable, which influences what opcodes are accessible (and their arguments and addressing modes), but is not the only constraint created. For example, in MOV-only compilation, the program's header does not need to consist only of MOV instructions, and so the compiler's output does not suffer the same severe code and data limitations that DOS EXEs do. (The executables it produces are extremely large and slow; they also seem to have non-MOV initialization code.) The MOV instruction is also very rich, and no versions of it are printable!

Of course, everyone knows that even unary numbers (just like one symbol repeated a given number of times) is Turing complete, via Godel encoding. So what's the big deal?

** 32. Acknowledgements **

The author would like to thank the fastidious SIBOVIK "Program" Committee for "Evaluating" my paper.

** 33. Bibliography **

[KNPH'14] Tom Murphy VII. "New results in k/n Power-Hours." SIGBOVIK, April 2014.

[MTMC'08] Tom Murphy VII. "Modal Types for Mobile Code." Ph.D. thesis, Carnegie Mellon University, January 2008. Technical report CMU-CS-08-126.

[LLVM'04] Chris Lattner and Vikram Avde. "LLVM: A Compilation Framework for Lifelong Program Analysis and Transformation." CGO, March 2004.

[CKIT'00] David Ladd, Satish Chandra, Michael Siff, Nevin Heintze, Dino Oliva, and Dave MacQueen. "Ckit: A front end for C in SML." March 2000. <http://smlnj.org/doc/ckit/>

[INTC'01] Intel Corporation. "IA-32 Intel Architecture Software Developer's Manual. Volume 2: Instruction Set Reference." 2001.

[ABC'05] Steve Mansfield. "How to interpret abc music notation." 2005.

[MOV'13] Stephen Dolan. "mov is Turing-complete". 2013.

[MVF'16] Chris Domas. "M/o/Vfuscator2". August 2015. <https://github.com/xoreaxeax/movfuscator>

Please see <http://tom7.org/abc> for supplemental material.



Figure 7. Printable X86

** Appendix **

Here is a histogram of every character that appears in this file. There are no non-printable bytes.

char	byte	number of occurrences	
	0x20	186844	
!	0x21	1042	
"	0x22	1616	
#	0x23	3216	
\$	0x24	344	This
%	0x25	3859	column
&	0x26	73	
'	0x27	312	
(0x28	8683	is
)	0x29	822	
*	0x2A	512	unintentionally
+	0x2B	193	
,	0x2C	1718	left
-	0x2D	27759	
.	0x2E	7325	blank.
/	0x2F	224	
0	0x30	1015	
1	0x31	1383	
2	0x32	542	
3	0x33	1110	
4	0x34	2231	
5	0x35	360	
6	0x36	299	
7	0x37	172	
8	0x38	295	
9	0x39	101	
:	0x3A	1418	
;	0x3B	270	
<	0x3C	402	
=	0x3D	985	
>	0x3E	34	
?	0x3F	94	
@	0x40	11820	
A	0x41	509	
B	0x42	320	
C	0x43	875	
D	0x44	1297	
E	0x45	2326	
F	0x46	715	
G	0x47	214	
H	0x48	249	
I	0x49	418	
J	0x4A	107	
K	0x4B	342	
L	0x4C	343	
M	0x4D	448	
N	0x4E	668	
O	0x4F	248	
P	0x50	2350	OR IS IT !?!
Q	0x51	8240	
R	0x52	143	
S	0x53	371	
T	0x54	358	
U	0x55	151	
V	0x56	82	
W	0x57	123	
X	0x58	2369	
Y	0x59	62	
Z	0x5A	136	
[0x5B	103	
\	0x5C	136	
]	0x5D	353	
^	0x5E	703	
_	0x5F	23595	
	0x60	19	
a	0x61	5196	
b	0x62	1221	
c	0x63	2364	
d	0x64	2294	
e	0x65	8596	
f	0x66	1340	
g	0x67	4248	
h	0x68	3251	
i	0x69	5307	
j	0x6A	1378	
k	0x6B	487	
l	0x6C	2981	
m	0x6D	1969	
n	0x6E	4833	
o	0x6F	5094	
p	0x70	1770	
q	0x71	268	
r	0x72	4313	
s	0x73	4849	
t	0x74	7157	
u	0x75	3332	
v	0x76	676	
w	0x77	1211	
x	0x78	836	
y	0x79	1023	
z	0x7A	172	
{	0x7B	46	
	0x7C	109	
}	0x7D	861	
~	0x7E	16972	
total		409600	

The following characters were inserted to make the above converge: 4853



Insect track

Debugging



25 Amazon Web Services: Field observations related to arachnid cohabitation

Riva Riley

26 Blackberry Debugging

Allison M. Gardner and Kristen S. Gardner

Title: Amazon Web Services: field observations related to arachnid cohabitation

Author: Riva Riley (a human person)

Key words/phrases: wolf spider, Amazon rainforest, arachnid appreciation, predator/prey interactions

Abstract: Humans and arachnids have had a long, tense coexistence that stretches out over the entirety of the humans' relatively puny evolutionary history. Behavioral studies even suggest that humans have an innate fear of arachnid locomotion, and it has been demonstrated many times that humans recoil instinctively when they see a spider scuttling along, minding its own business. Here I present field observations made in the Peruvian Amazon Rainforest, Madre de Dios locality. My field experience provides evidence that the instinctive response humans feel toward arachnids may be reversed, and that an explicit symbiosis can form between humans encroaching on wild territory and the spiders that colonize their easily infiltrated dwellings. This symbiosis provides the spider with easily navigable terrain and a steady supply of prey (mostly insects drawn by the crumbs and refuse of human lifestyles), and provides to the human a cleaner environment without the annoyance of pests like roaches and the potential harm of disease vectors, like mosquitoes.

Experimental subjects: Unwilling in both cases, the subjects of this case study were the author, a field scientist in the Amazon rainforest, and an adult wolf spider the author called Octavio starting from day 11 of their cohabitation.

Catalogue of interactions:

Week one: I discovered the den of the 20cm long (torso ~7.5cm longitudinally, legs 10mm in circumference and well-guarded with numerous bristles) wolf spider on the second evening after my arrival the Centro de Investigación y Capacitación Río Los Amigos (CICRA). It was three hours past sunset and the only light came from my small flashlight- I saw in the thin beam of my light that the hollowed out area behind my sink was crawling (literally) with a beheaded cockroach carcass. I slowly turned my head, and the 20cm wolf spider was whirling down my wall. I shrieked and lifted my legs off the ground- it was fortunate I was still on the toilet. My vocalizations did not seem to have any effect on the spider's behavior (though the other scientists did protest). The spider arrived behind the sink, grabbed the cockroach body, and began to eat. I was horrified. The following morning, the spider was still there, as it was the following night, and the morning after that.

Week two: I had been tiptoeing around a spider roughly the size of a Chihuahua for several days now. I was scared to use my own bathroom and began to lose hope that it would leave. I considered asking one of the other scientists to remove it for me, but I did not want to lose what we colloquially referred to as 'field biologist trail cred' on a spider (the social dynamics and hierarchies of field stations are the subject for another paper). I also considered using my heavy snake-proof boots to pummel it to death, but when I got a look at the hulking scaffold of spider flesh I was up against, I became concerned. If I failed to defeat the spider, I might start an escalating series of antagonistic interactions whose natural end I did not want to discover. I

began to use my bathroom carefully, noting the spider's habits. The spider was mostly nocturnal, and so I did not use the bathroom during the night for several days. Its den is kept mostly clean- it eats or removes stray insect parts. At the end of the week I discovered that singing a soft song and shining my light alerts the spider to my presence, which it has also decided to tolerate. As a response to those stimuli, the spider retreats to its den, and so I learn how to use my bathroom while sharing the territory. At the end of the week I named the creature 'Octavion'. The other field scientists all complained about cockroaches getting into their luggage; I had no idea what they are talking about.

Week three: Octavion and I have come to know one another. I have composed a specific song to sing when communicating with the spider, and he seemed to respond differently to this song as compared to ABBA's 'Dancing Queen'. I talked to him as I brushed my teeth and combed my hair, and he would cue in to the vibration patterns of my voice and sit on a ledge to listen. On the last day of this week I awoke in the middle of the night only to see Octavion's underbelly mere inches from my face- he had climbed onto my mosquito net. Disquieted but not terribly so, I fell back asleep, and from here on out, I receive notably fewer mosquito bites through the netting in the nighttime. Octavion consistently chooses not to bite me through the mosquito net.

Week four: A second wolf spider arrived in my room. I was not enthusiastic, as it sat on my hairbrush and could jump over a meter in a single bound. It was gone within a few hours, and Octavion spent longer than usual in its den. I wondered if Octavion were female, and how I would deal with baby spiders. I began thinking of ways to work out a system for inviting guests over. I have already alerted my friends at the field station to follow appropriate protocol if we are congregating in my room in the evening.

Week five: With no parting communication, Octavion has gone. His den is empty and I spotted a roach in my cabinet. This made sense, as he was a spider, and spiders do not communicate information like this. Nonetheless, I experienced a sense of loss. Another field scientist mentioned a large wolf spider lurking in a different bathroom. I hoped Octavion was happy without me, but not too happy.

Weeks six-eight: Octavion eventually leaves the station. I try not to worry about him, as he is a spider, and cannot worry about me.

Discussion:

My room had significantly fewer pests during the time of Octavion's cohabitation, including cockroaches, cicadas, crickets, and beetles.

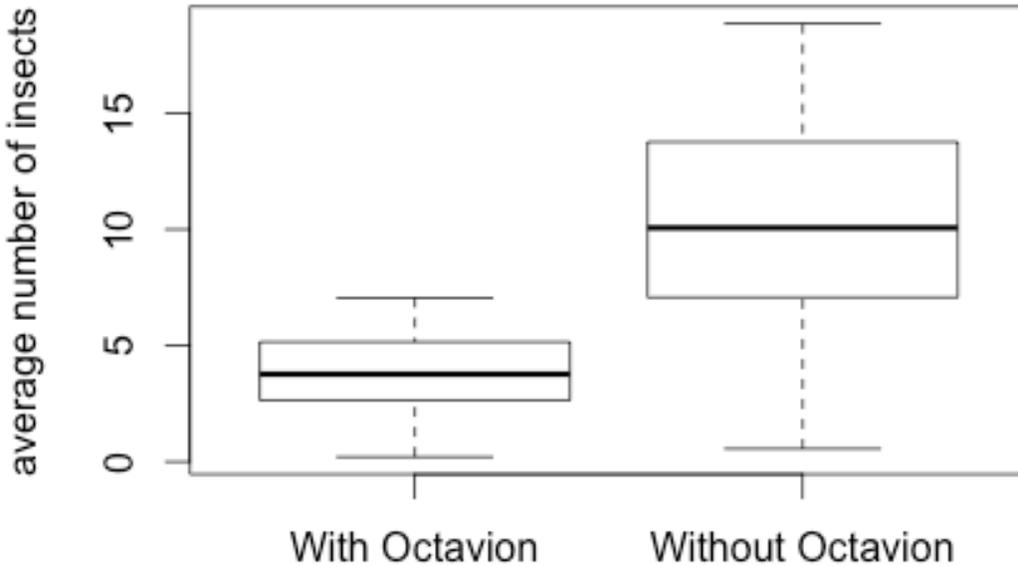


Figure 1: average number of insect pests in my jungle bedroom with Octavion and without Octavion ($p=0.02$)

Octavion was not messing around. My affection for him seems unnatural to other humans, and it requires effort and coaxing to get others to appreciate the great service he did for me. In fact, I wonder if my affection was not simply an adaptive response to help overcome humanity's aversive reaction to arachnids in the face of their overwhelming benefit. If other humans could tolerate our arachnid brethren more openly, fewer pests and insect bites might become the norm. I believe there is a path to more directed mutualisms between spiders and humans that might completely change life as we know it. Spiders are small and mobile, and can access areas humans cannot. We should take advantage of this tremendous resource literally crawling at our feet (don't jump!). By establishing communications between humans and arachnids, we may be able to harness their ick factor for our own good. These associations will minimize potentially dangerous insect bites and provide companions for the people they live with. To take advantage of an arachnid association, I intend to establish more formalized communication protocols to help cement the bond between human and arachnid in the modern era.

Blackberry Debugging

Allison M. Gardner^{1*} and Kristen S. Gardner^{2*}

¹School of Biology and Ecology, University of Maine, Orono, ME

²Department of Computer Science, Carnegie Mellon University, Pittsburgh, PA

*These authors did not contribute equally to this work.

Background. *Culex pipiens* Linnaeus (Diptera: Culicidae), an important mosquito vector for West Nile virus in urban landscapes throughout the northeastern and midwestern United States, oviposits in a variety of natural and artificial containers such as small ponds, discarded tires, and storm water catch basins. These habitats are mainly fueled by plant-based detritus from the surrounding terrestrial vegetation. Detritus type and quantity determine the composition and abundance of microbial communities that form in container habitats as microbes break down terrestrial leaf litter. In turn, these bacteria and fungi provide a direct food source for mosquito larvae and influence oviposition behavior of gravid female mosquitoes through emission of oviposition attractants and stimulants. Thus, detritus from terrestrial plants and its associated microbes play a critical role in determining vector distribution, relative abundance, and life history traits that are important for vector-borne pathogen transmission including adult body size, longevity, biting rates, and vector competence. In this study, we test the hypotheses that leaf detritus of three native and three invasive shrubs asymmetrically affects oviposition site selection and adult emergence rates of *Cx. pipiens*.

Methods. Six focal plant species were selected among shrubs common within the geographic range of *Cx. pipiens*: *Lonicera maackii* (Dipsacales: Caprifoliaceae; Amur honeysuckle), *Elaeagnus umbellata* (Rosales: Elaeagnaceae; autumn olive), *Rosa multiflora* (Rosales: Rosaceae; multiflora rose), *Rubus allegheniensis* (Rosales: Rosaceae; blackberry), *Sambucus canadensis* (Dipsacales: Adoxaceae; elderberry), and *Amelanchier laevis* (Rosales: Rosaceae; serviceberry).

To test the hypothesis that leaf detritus species in the aquatic environment affects oviposition site selection of *Cx. pipiens*, six oviposition traps each containing 4 L of tap water and 80 g of fresh whole leaves of one of the six shrub species were placed 1 m apart from each other in partial shade at five sites located within a 5 km radius in a residential neighborhood. The 30 oviposition traps were monitored for egg rafts daily and the number of egg rafts collected in each substrate from June 24 to August 5, 2013 was recorded. A general linear mixed model (GLMM) with repeated measures was used to compare the abundance of egg rafts collected by leaf substrate, day, and their interaction throughout the study period.

To test the hypothesis that leaf detritus species affects adult emergence rates of *Cx. pipiens*, larvae were obtained by collecting egg rafts from five sites using grass infusion-baited oviposition traps. Egg rafts were individually hatched in petri dishes containing deionized water. To test for the effect of leaf substrate on intraspecific competition, 18 treatments were established with five replicates per treatment. Each treatment included one of three densities of first instar larvae of *Cx. pipiens* (10, 20, or 40 per container) and 360 mL infusion of one of the six leaf detritus species in 400 mL tri-pour beakers. Infusions were prepared by fermenting 80 g of fresh leaves of each plant species in 4 L of tap water for 7 days. The containers were monitored daily and pupae were removed from containers and housed individually in cotton-sealed plastic vials with deionized water. A GLMM with a factorial treatment structure was used to test the fixed effects of leaf species, competition, and their interaction on *Cx. pipiens* emergence rates.

Results. The number of egg rafts laid in oviposition traps containing the leaves of different native and invasive shrubs varied within and among leaf detritus species over the collection period, with significant

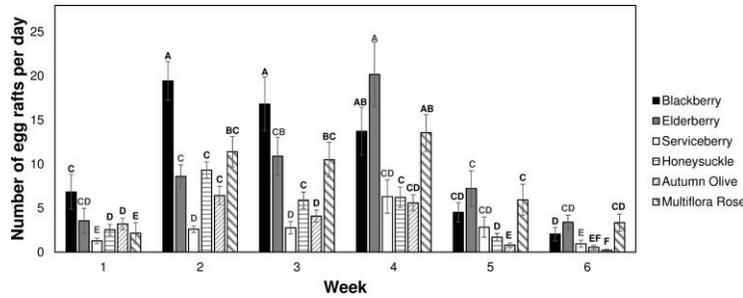


Figure 1. Mean (± 1 standard error) for *Culex pipiens* egg rafts collected in oviposition traps per day from June 24 to August 5, 2013 (6 weeks) by leaf detritus treatment. Letters indicate significant pairwise differences at $\alpha = 0.05$.

Mosquito emergence rates varied across leaf detritus types and larval densities, with a significant interaction between leaf species and density ($F = 7.33$; $df = 10, 72$; $P < 0.01$; Figure 2). The lowest emergence rates were observed in blackberry and multiflora rose infusion; no adults emerged in the latter leaf detritus species at any larval density. The highest emergence rates were observed in honeysuckle and autumn olive infusions, although autumn olive-reared mosquitoes experienced a significant decline in emergence at the highest density while honeysuckle infusion mitigated the deleterious effects of intraspecific competition even at high larval densities. Among all other leaf species except for elderberry, higher larval densities yielded significantly lower emergence rates than lower larval densities.

Discussion. Using a combination of laboratory and field experiments, we identified two invasive shrubs that may promote growth and emergence of an important mosquito vector relative to native shrub species by improving the nutritional quality of the larval environment via leaf detritus inputs. *Culex pipiens* emergence rates were significantly higher in leaf infusions of honeysuckle and autumn olive compared to the other shrub species. We also noted that the deleterious effects of intraspecific larval competition were mitigated in honeysuckle treatments. These results complement a growing body of field studies that suggest landscaping with exotic – and potentially invasive – plants has the potential to influence local larval and adult mosquito abundance and distribution.

Our comparisons of native and invasive leaf detritus species on larval development and oviposition facilitated the discovery of a naturally-occurring ecological trap for *Cx. pipiens*. Ecological traps occur due to a mismatch between the attractiveness of a habitat and its quality for reproduction. The greatest number of egg rafts was collected in water containing leaves of blackberry, a native plant species found throughout the geographic range of *Cx. pipiens*. However, in laboratory assays, exceptionally low mosquito emergence rates were observed in blackberry infusions, with fewer than 20 percent of larvae surviving to eclosion even at the lowest larval density. Infusion of multiflora rose, an exotic shrub of limited importance in the Midwest but highly invasive in the northeastern United States, similarly

effects of leaf species ($F = 7.25$; $df = 5, 20$; $P < 0.01$) and day ($F = 23.70$; $df = 39, 912$; $P < 0.01$) but not their interaction (Figure 1). Throughout the experiment, the greatest number of egg rafts were collected in blackberry and elderberry leaf infusion, the lowest number of egg rafts per day were collected from water containing serviceberry, autumn olive, and honeysuckle leaves, and an intermediate number of egg rafts were collected from water containing multiflora rose.

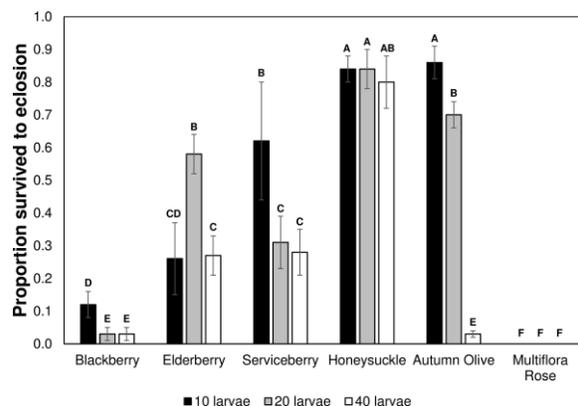
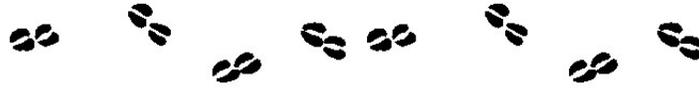


Figure 2. Mean (± 1 standard error) for *Culex pipiens* male and female emergence rates across intraspecific competition by leaf detritus treatments. Letters indicate significant pairwise differences at $\alpha = 0.05$.

yielded lower emergence with no mosquitoes developing to eclosion across all density treatments. However, gravid females were better able to discriminate against this leaf detritus species and consequently water containing multiflora rose leaves collected fewer egg rafts than water containing blackberry leaves. Future research will determine whether exploitation of this ecological trap may yield a novel “attract-kill” approach (i.e., Blackberry Debugging) to control mosquito larvae in closed aquatic environments, such as rain barrels, buckets, and storm water catch basins.

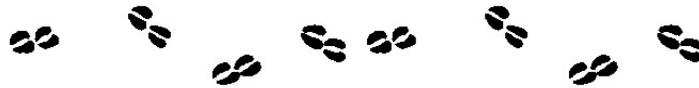
In summary, we observed elevated emergence rates and more rapid development among *Cx. pipiens* mosquitoes reared in infusions of honeysuckle and autumn olive leaves, two exotic, invasive shrubs that occur throughout much of the northeastern and midwestern United States. In contrast, we discovered mosquito emergence was significantly reduced among mosquitoes reared in infusions of native blackberry and exotic multiflora rose leaves compared to those exposed to other leaf detritus species. Our results have applications in two areas. First, our findings that some exotic, invasive shrubs are favorable for mosquito production may be relevant to mosquito control and invasive plant management in the range of *Cx. pipiens*. Second, our discovery of a previously unknown ecological trap for an important vector of West Nile virus has the potential to lead to novel alternatives to conventional insecticides in mosquito control, exploiting the apparent “attract-kill” properties of this native plant species.

Acknowledgments. AMG’s former PhD advisors think this work only has been published in *Parasites and Vectors* (doi: 10.1186/s13071-015-0941-z). Let’s keep it that way.



Moose track

Impure Math and \Big Data



27 Fake news logic

Will Nalls

28 RRR for UUU: Exact analysis of pee queue systems with perfect urinal etiquette

Kristen Gardner and Ziv Scully

29 The next 700 type systems

Carlo Angiuli

30 A modular approach to state-of-the-art big data visualization

Keith A. Maki

31 Efficient computation of an optimal portmantout

David Renshaw

FAKE NEWS LOGIC

WILL NALLS
CMU PHILOSOPHY
MARCH 2017

ABSTRACT. In this paper, we attempt to combat some of the confusion surrounding FAKE NEWS by providing a formal logical framework to encode and evaluate statements in a consistent manner. We expand the typical language for epistemic logic language with two novel modalities, FN and AF , and provide semantics for both. We demonstrate that this language stands in accord with typical instances that we are trying to capture, and model the tragic *Bowling Green Massacre*. Lastly, we briefly indicate how one should think about axiomatizing this logic.

1. INTRODUCTION

In late 2016, the problem of FAKE NEWS was brought to the attention of the public by the *corrupt media*.¹ The *corrupt media* claimed wildly and without base that several forms of social media had witnessed the arrival of a multitude of sites publishing FAKE NEWS for a profit. Most thankfully, the record was set straight with the arrival of new executive leadership, henceforth referred to as the d. The d rescued the term and restored its status as a descriptor of all major news outlets – that is, the *corrupt media*. However, many have most unfortunately misinterpreted his decrees regarding FAKE NEWS. In this paper, we address this misinterpretation by formalizing his commentary in a modal logic, thus dispelling the provably false claim that the d was, or ever could be, wrong.

We expand the language of epistemic logic, \mathcal{L}_{EL} , to the language of FAKE NEWS, \mathcal{L}_{FN} . We motivate the informal interpretation of the expansions with important phenomena, and distill these interpretations into a formal semantics for what we call *FAKE NEWS logic* (FNL). We apply these semantics to model a well-known scenario, and consider what might be involved in designing a sound and complete logic for the language.

2. LANGUAGE AND SEMANTICS: IDENTIFYING FAKE NEWS

We begin with standard epistemic logic, which takes as given a countable set of propositions, Φ , and a finite set of agents, G . Since we are only concerned with the knowledge statements of one individual, the d, G will be the singleton $\{d\}$. The language of epistemic logic, \mathcal{L}_{EL} , is generated inductively:

¹The term was first brought to the fore by Mark Zuckerberg in an inspirationally unempathetic remark in response to the problem of pervasive misinformation. See <https://www.forbes.com/sites/kalevleetaru/2017/02/17/did-facebooks-mark-zuckerberg-coin-the-phrase-fake-news>.

$$p \in \Phi \mid \neg\varphi \mid \varphi \wedge \psi \mid K_d\varphi$$

Recall that formulas are evaluated in models at worlds. A model $M = \langle W, \sim_d, V \rangle$ consists of the following:

- A set of possible worlds, W .
- An epistemic relation over these worlds, $\sim_d \subseteq W \times W$. $(w, w') \in \sim_d$, sometimes writted as $w \sim_d w'$, should be read as ‘from w , the d has epistemic access to w' ’.
- An evaluation function $V : \Phi \rightarrow 2^W$, which assigns to every proposition p a set of worlds JpK_d where the proposition is true. This evaluation will be extended in a coherent way to the entire language via the semantics given below.

p is true at w in M just when $w \in [p]_d$, written $M, w \models p$.² We note immediately that standard semantics for the K operator will not suffice for present considerations. Typically, $K_d\varphi$ is true in a pointed model M, w when for every other world, w' , epistemically accessible by d from w , it is the case that $M, w' \models \varphi$. However, it occurs frequently that the assertion $K_d\varphi$ is handed down from above when it is commonly known that at *no* possible world is φ true. Thus, maintaining typical semantics for the K operator would require that we render the epistemic access relation, \sim_d , empty; we disregard this possibility and offer the alternative semantics for K_d .³

$$(M, w) \models K_d\varphi \text{ iff for some } w' \text{ in } M, w \sim_d w' \text{ and } (M, w') \models \varphi$$

Note that this is simply the dual of the typical semantics for K_d , read as, ‘the d knows that φ is true if d has access to at least one world where φ is true.’ We find this reading of the K operator to be a faithful translation of the statements under consideration.

Importantly, this reading of the K operator allows for the d ’s frequently contradictory proclamations of knowledge; were we to preserve the semantics of the K operator, if \sim_d were nonempty then our logic would reduce to triviality.⁴⁵

We write $[\varphi]_V$ for any $\varphi \in \mathcal{L}_{EL}$ to denote the set of worlds at which φ is true according to V and the given semantics.

It will be worthwhile presently to identify some of the properties of \sim_d . Firstly, it is clear that \sim_d is not reflexive, as no evidence guarantees that the d has any access to the actual world. Indeed, since all utterances of the d seem to engage with a different reality, we characterize \sim_d as *anti-reflexive*: $(\forall w \in W)(w \not\sim_d w)$.

²We endorse a typical abuse of notation by saying ‘ w in M ’ instead of ‘ w in W ’.

³If \sim_d were empty, this would mean that the d has access to *no* possible worlds. Clearly, this is an interesting and plausible possibility that should be explored; we disregard this possibility here because the tools offered by epistemic logic are not suitable to such an investigation.

⁴Again, this is a possibility which merits further investigation – indeed, there is much reason to suspect that any formal characterization of the d ’s language should reduce to triviality – but we set this aside for another discussion.

⁵For an incomplete listing of contradictions, refer to <http://www.politico.com/magazine/story/2016/05/donald-trump-2016-contradictions-213869>.

This is supported by several decrees from the d. We posit furthermore that the d has no *extended access* to the actual world, either. That is, there is no chain of worlds w, w', w'', \dots such that $w \sim_d w' \sim_d w'' \sim_d \dots \sim_d w$. Thus, we stipulate that \sim_d is a *acyclic*.

We now turn to our additions to the language, the modal operators FN and AF . We begin with FN . FN should, intuitively, apply only to formulas which are regarded as FAKE NEWS from the present world. Furthermore, we claim that FN should apply only to *propositions*: what the d knows or doesn't know is not subject to discussion – especially not by the *corrupt media*.⁶ While FN will be designed to apply only to propositions, we may informally regard it as applying to boolean combinations of propositions, too.

We first note that FN stands in a close relationship with the modality K_d . In particular, if the d does not know p , then p is FAKE NEWS. Formally, this is stated as follows

$$(M, w) \models \hat{K}_d \neg p \text{ implies that } (M, w) \models FNp$$

The novel component of FAKE NEWS, however, is that *it includes anything disliked by the d*. Formalizing this essential information will require a minimality condition on the set of propositions, Φ .⁷

For every proposition p in Φ , there is the proposition $l_p \in \Phi$

We read l_p as ‘the d likes p’. The valuation function V now acts on this expanded set; $M, w \models l_p \wedge p$ denotes, for instance, that p is true at w and the d likes that. The additional sufficient condition for FN may be informally stated as follows: if the d does not like p and has the conceptual capacity to regard p as an active possibility, then p is FAKE NEWS. Formally:

$$(M, w) \models \neg l_p \wedge K_d p \text{ implies } (M, w) \models FNp$$

These are the only conditions under which FNp will be true at a world. Note that when taken together, these conditions imply that the necessary and sufficient condition for FNp is just $\neg l_p$ – as desired, p is FAKE NEWS just when the d does not like p .

We now turn to the second notion we represent in our language, that of *alternative facts*. This concept was first introduced by one of the d's top henchpersons in defense of a proclamation made by one of the d's meeker and less articulate henchpersons: “You're saying that it's a falsehood and... [we] gave alternative facts to that.”⁸ We interpret this comment to point to the d's metaphysical take on truth: the d – and by extension, his underlings – subscribes to the belief that truth is a purely linguistic artifact; that there are no external and independent facts to which true utterances correspond. Alternative facts

⁶Indeed, it has been observed that even the d's underlings are hesitant to make any declaration regarding the d's epistemic state.

⁷Note that this minimality condition does not increase the cardinality of Φ .

⁸Witness the exchange at <http://www.nbcnews.com/meet-the-press/video/conway-press-secretary-gave-alternative-facts-860142147643>.

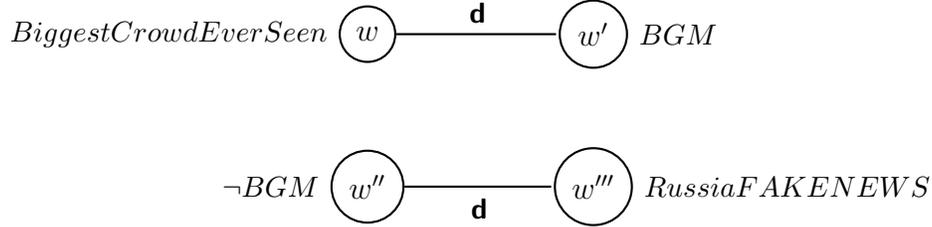
are introduced, then, to capture this flexibility of truth – any statement which *could* be regarded as truthful according to some linguistic scheme may be regarded as an alternative fact. We introduce the modality AF with the following semantics:

$$M, w \models AF\varphi \text{ iff there is some } M', w' \text{ such that } M', w' \models \varphi$$

In words, $AF\varphi$ holds at a world if there is some model and some world that satisfy φ ; or, φ is not a contradiction. The permissive nature of this concept captures the sort of semantic freedom exercised by the d. Note that alternative facts are independent of worlds and models; their assertion relates in no way to *local* circumstances.

We name the language of epistemic logic, \mathcal{L}_{EL} expanded with the operators FN and AF , the language of FAKE NEWS logic, \mathcal{L}_{FN} . To demonstrate its utility, we capture a well-known phenomenon relating to the matters of FAKE NEWS.

The tragic *Bowling Green Massacre* was despicably underreported by the *corrupt media*, until brought to light by the same top henchperson who introduced alternative facts.⁹ The national concern over the matter was captured with the tweet, “The real lesson from Bowling Green. Who will cover? Who will care?”¹⁰ We model the situation of the Bowling Green Massacre (BGM) as follows:



We leave the interpretation of this model (in particular, the undirected edges) up to the reader.

In closing, we provide some considerations towards the construction of a sound and complete logic for this language. Clearly, the K operator may be easily axiomatized. The FN operator may also be easily axiomatized, if a relation is introduced to the language to capture the correspondence between propositions p and l_p . However, the AF operator is difficult to capture, as its truth conditions include metatheoretic requirements. The difficulty is in importing such requirements into a syntactic setting, without significantly strengthening the language. We propose that this particular difficulty is not a mark *against* our characterization, however. Indeed, we think that any faithful description of these phenomena will render impossible any kind of coherent corresponding derivation system; such a system would undermine the very flexibility of truth which is so essential to the d’s public image.

⁹We send our thoughts and prayers to those whose attempts to stifle cultural diversity were thwarted by the unfair and baseless allegations that the Bowling Green Massacre never occurred.

¹⁰This tweet may be found at <https://twitter.com/KellyannePolls/status/827583711126360065>.



CONFIDENTIAL COMMITTEE MATERIALS

SIGBOVIK 2017 Paper Review

Paper 91: Fake News Logic

Reviewer D

Ratings: Absolutely great ratings, HUGE audience.

Confidence: I guarantee you there's no problem. Believe me.

The Philosophy Department's got great people. They love me over there. It's no CS department, but boy they've got a lot of computers. My nephew would love it over there. Will's a great guy, known him a long time, trustworthy, not like the media, believe me.

The global elite write a lot of papers on modal logic. They tell us how to think. Look what happened the last time we used one of those epistemic logics. Have you even seen their Kripke models? Their Kripke models remind of Rosie O'Donnell, 2/10, tops. Just sad. Your Kripke models are a real 10, just like my daughter. Great models.

I don't like all this talk about possible worlds, though. We're getting killed at global trade, killed by Mexico, killed by China, killed by NATO. Let's stop spending money on all these other possible worlds. Wasted money. Let's spend that money on JOBS.

I know how to generalize this result to the syntactic setting. I know a lot about syntax. I've got the best syntax. Wharton grad. Make \mathcal{L}_{FN} strong. Lots of people are saying S_5 doesn't have a derivation system anyway. They say we don't have a derivation system but then you go ask them "Hey where's your derivation system" and they don't have one! Despicable.

RRR for UUU: Exact Analysis of Pee Queue Systems with Perfect Urinal Etiquette

Kristen Gardner
Computer Science Department
Carnegie Mellon University
ksgardne@cs.cmu.edu

Ziv Scully
Computer Science Department
Carnegie Mellon University
zscully@cs.cmu.edu

ABSTRACT

Queueing systems with multiple servers operating in parallel, such as the $M/M/k$ model, have been extensively studied. However, most prior literature examines the limited case in which that all servers may operate simultaneously. This is despite the fact that many practical queueing systems encounter constraints that make this an unrealistic assumption. In this work, we investigate one such setting: the men’s lavatory, in which a strict etiquette requires that no two adjacent urinals be in use at the same time. We introduce and analyze a new queueing model, the Context-2 Unease Processing Network (C2UPN), which formalizes a row of urinals used with perfect etiquette. We derive exact results for a row of 3 urinals (UUU) using the Recursive Renewal Reward (RRR) technique. Remarkably, our method generalizes to many other urinal topologies, including longer rows and cyclic configurations.

ACM Reference format:

Kristen Gardner and Ziv Scully. 2017. RRR for UUU: Exact Analysis of Pee Queue Systems with Perfect Urinal Etiquette. In *Proceedings of SIGBOVIK 2017, Pittsburgh, PA, USA, March 31, 2017 (SIGBOVIK ’17)*, 5 pages.

1 INTRODUCTION

Imagine you are at a conference, and the second coffee break is approaching. Most attendees are filled with the free diuretics provided during the first coffee break and have only one question on their minds: how long will they have to wait in line to use the restroom once the break starts? It is of utmost importance to sustain only a short queueing time. Despite the universality of this problem, there is shockingly little theoretical work analyzing response time in bathrooms. Instead, most prior work on bathrooms focuses on anthropological studies of human behavior in bathrooms.

In this paper, we turn to queueing theory to present new exact analysis of response time in one specific category of bathroom queueing problem: the urinals-only setting. Here customers arrive to the system, wait in the queue for an available urinal, and depart the system upon completing their urination. While many bathrooms consist of both urinals and stalls, we choose to focus on the urinals-only setting for two reasons. First, often in bathrooms that offer both urinals and stalls, people opt to only use the stalls even if there are vacant urinals [2]. Second, the introduction of stalls necessitates the existence of both “type-1” and “type-2” customers that have different service times; this complicates the analysis.

Urinal queueing exhibits several unique properties not present in other bathroom queueing settings. Most notably, here we must consider the degree of *urinal etiquette* exhibited by urinators. Urinal

etiquette has to do with the number of vacant urinals left between urinators. The three possible urinal etiquette degrees are:

- (1) No etiquette. Here an arriving urinator will use any vacant urinal, regardless of whether the adjacent urinals are vacant or occupied.
- (2) Partial etiquette. Here an arriving urinator may choose to use a vacant urinal that is adjacent to an occupied urinal (for example, if the urinator is experiencing a high degree of urgency). The urinator may also choose to join the queue if the only vacant urinals are adjacent to occupied urinals.
- (3) Perfect etiquette. Here an arriving urinator *never* uses a vacant urinal that is adjacent to an occupied urinal.

In all three settings, we assume that an arriving urinator will always use a vacant urinal that is not adjacent to an occupied urinal. We note that Justus argues that the “buffer zone” is mandatory, meaning that an arriving urinator may never use a vacant urinal that is next to an occupied urinal, except in cases of unusually high load [5]. We agree, hence in this paper we focus on the perfect etiquette setting.

Our main contribution in this paper is the first exact analysis of response time in urinal systems. We begin by studying 3- and 5-urinal systems configured in a standard topology in which the urinals are arranged in a straight line. Our approach involves modeling the system using a Markov chain and apply the Recursive Renewal Reward (RRR) technique to solve the chain exactly. In Section 4 we consider alternative urinal topologies and investigate the conditions under which our approach allows us to develop exact analysis. Finally in Section 5 we discuss directions for future work.

1.1 Related Work

As noted above, most of the related work focuses on human behavior. Cahill et al. conduct an extensive observational study of humans in bathrooms and find that while unacquainted individuals typically avoid conversation at urinals, people who already know each other often converse while urinating (though they avoid making eye contact) [2].¹ The authors also point out that closed stalls make excellent hiding places while conducting observational studies in bathrooms [2]. Empirical results also indicate that women spend a significantly longer time using bathrooms than men do [1]. Perhaps this is because 85% of women choose to “crouch” instead of sitting directly on public toilet seats, which can reduce the average flow rate by 21% [8]. We refer the reader to [9] for a survey of

¹Our personal sense of urinal etiquette, finely tuned by more than two decades of combined urinal experience, discourages urinal conversation, particularly between students and their advisors.

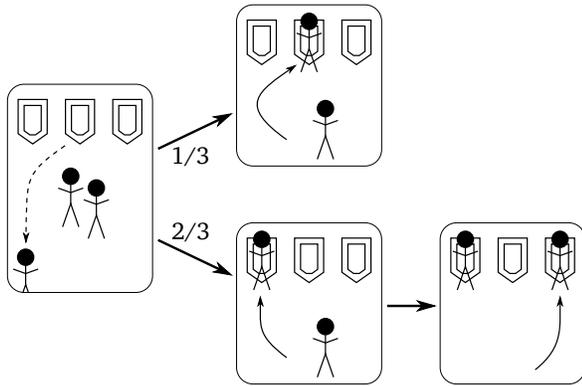


Figure 2.1: The urinal selection process with 3 urinals in a row (viz. the $M/M/3/C2UPN$ queue). A single completion at the center urinal leaves the next urinator with a choice between 3 available urinals. If they choose an edge urinal, then a second urinator can also enter service.

other related behavioral results, perhaps to read while using the bathroom (69% of people use their phones while on the toilet [3]).

All of the above work, though interesting, is orthogonal to our mathematical approach to the urinal problem. To our knowledge, the only existing theoretical work on urinal usage is [6]. The paper considers a setting in which a urinator enters the system and needs to choose which urinal to occupy so as to maximize his privacy, i.e., the time until an adjacent urinal becomes occupied. Unfortunately the model makes several uncomfortable assumptions, including that the urination duration is infinite, so urinators never leave the bathroom. Our work allows for finite urination duration and assumes perfect etiquette, so privacy is always maintained.

2 SYSTEM MODEL

We model *urinators* as servers of fixed service rate 1. Lavatory users, or *urinators*, have i.i.d. service requirements drawn from a specified distribution and arrive according to a specified stochastic process. There is a queue of infinite capacity holding waiting urinators, who enter service when possible in first-come, first-served order.

In traditional queueing, all servers are available at all times. Unfortunately, life is not so simple in the men’s lavatory. In practice, adjacent urinals cannot both serve urinators at the same time. We capture this as an *unease graph*: vertices represent urinals, and edges represent pairs of urinals that cannot be occupied simultaneously. While a urinal is serving a urinator, its neighbors in the unease graph become unavailable until service at the urinal completes.

The urinator waiting at the head of the queue enters service at the first urinal to become available. When multiple urinals to become available simultaneously, the urinator occupies a urinal chosen uniformly at random from the set of available urinals. Note that a single completion can enable more than one waiting urinator to enter service. Figure 2.1 illustrates an example of the urinal selection process.

In the vast majority of men’s lavatories, the urinals are assembled in a small number of rows. In this case, the unease graph is a union of paths, with edges between adjacent urinals in the same row. We call a single row of urinals a *Context-2 Unease Processing Network* (C2UPN), as each occupied urinal makes up to 2 urinals unavailable.

For the remainder of this paper, we consider the Markovian case, which has exponential service and interarrival times. Urinator service requirements are distributed exponentially with rate μ , and the lavatory experiences a Poisson arrival stream of urinators with rate λ . In Kendall notation, this is the $M/M/k/C2UPN$ queue.

A natural generalization of the C2UPN is to consider other urinal topologies arising from unease graphs beyond paths. Because the number of urinals made uneasy by an may be an arbitrary number N , we refer to this as the *Context- N Unease Processing Network* (CNUPN). We discuss some CNUPN systems in Section 4.

3 ANALYZING MARKOVIAN C2UPN QUEUES

Our chief weapon for analyzing $M/M/k/C2UPN$ systems is *Recursive Renewal Reward* (RRR), a technique pioneered by Gandhi et al. [4]. Originally introduced to analyze the $M/M/k$ queue with exponentially distributed setup costs, RRR enables exact analysis of many queueing systems with Markov chains that are infinite in just one dimension. Roughly speaking, we can think of a one-dimensionally infinite Markov chain as consisting of several “layers” of states. If the chain is eventually periodic and transitions between layers in the repeating region are all one-way, we can apply RRR to compute variety of metrics. Specifically, we find the z -transform of the number of jobs in the queue, from which various useful metrics, such as expected system response time, follow easily.

3.1 Summary of RRR

Here we give a high-level overview of RRR, referring the reader to Gandhi et al. [4] for a more detailed exposition. RRR solves the following problem: given a Markov chain with certain nice properties, find the average value of some time-varying *reward rate*, which depends only on the current state. For example, to find the average number of jobs in the queue, $E[N]$, we set the reward rate in each state to the number of urinators in the queue. (We write N instead of the usual N_Q for the number in queue to reduce clutter.)

We designate some set of states as *home states*. For queueing systems, there is usually a single home state, namely the empty system. Given home states, a *renewal cycle* is the time interval between transitions into the set home states. A classic result of renewal-reward theory tells us that

$$E[\text{reward rate}] = \frac{E[\text{reward accumulated in a renewal cycle}]}{E[\text{length of a renewal cycle}]}.$$

For certain Markov chains, such as that of the $M/M/3/C2UPN$, the two quantities right-hand side are very easy to compute!

It is clear how to use RRR to compute expectations. To compute more complicated metrics, such as $\text{Var}(N)$, we just use a more complicated reward rate function. In particular, all moments of N can be computed from its z -transform, $\tilde{N}(z)$. Because $\tilde{N}(z) = E[z^N]$, we can compute it by setting the reward rate of each state to be z^n , where n is the number of jobs in the queue when the chain is in that state.

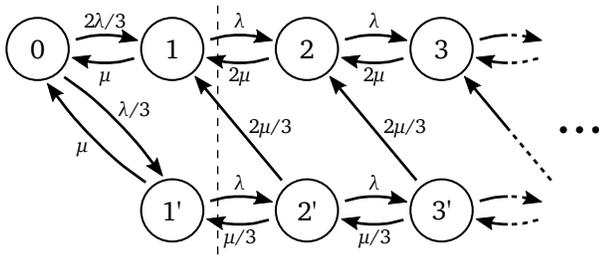


Figure 3.1: The M/M/3/C2UPN CTMC. The repeating portion is to the right of the dotted line, starting at 2 and 2'.

3.2 Exact Analysis of the M/M/3/C2UPN Queue

The continuous time Markov chain (CTMC) for the M/M/3/C2UPN queue is shown in Figure 3.1. The chain has two layers: a “sad” layer, in which the middle urinal is occupied, and a “happy” layer, in which the middle urinal is idle. We label the happy and sad states with n total urinators in the system with n and n' , respectively. (There is no state $0'$ because the middle urinal cannot be occupied if there are no urinators.)

Because our model is in continuous time, the probability of two urinators completing at the exact same time is 0. This means that in the repeating portion of the chain, it is possible to go from sad to happy (see Figure 2.1) but not vice versa, because a transition from happy to sad would require simultaneous departures to make the center urinal available. The only transition into the sad layer is from state 0, when a urinator arriving at an empty system occupies the center urinal.

We now apply RRR to find $\hat{N}(z)$. State 0 serves as our single home state. The first step is to compute the expected renewal cycle length, which we write as L . We will need the following definitions.

- Let L_i for $i \geq 1$ be the expected amount of time it takes to go left from i , possibly visiting states to the right in the meantime. Going left from i always ends up at $i - 1$.
- Let L'_i for $i \geq 1$ be the expected amount of time it takes to go left from i' . Going left from i' may end up at either $i - 1$ or $(i - 1)'$.
- Let q be the probability that going left from i' lands at $i - 1$, and let $q' = 1 - q$ be the probability that going left from i' lands at $(i - 1)'$.

Examining the Markov chain, we see that

$$\begin{aligned} L &= \frac{1}{\lambda} + \frac{2}{3}L_1 + \frac{1}{3}L'_1 \\ L_1 &= \frac{1}{\mu + \lambda} + \frac{\lambda}{\mu + \lambda}(L_2 + L_1) \\ L'_1 &= \frac{1}{\mu + \lambda} + \frac{\lambda}{\mu + \lambda}(L'_2 + qL_1 + q'L'_1). \end{aligned}$$

Because the Markov chain repeats after 2 and 2', we know $L_3 = L_2$ and $L'_3 = L'_2$, giving us

$$\begin{aligned} L_2 &= \frac{1}{2\mu + \lambda} + \frac{\lambda}{2\mu + \lambda}(2L_2) \\ L'_2 &= \frac{1}{\mu + \lambda} + \frac{\lambda}{\mu + \lambda}((1 + q')L'_2 + (1 - q')L_2). \end{aligned}$$

Given q' , these linear equations are easily solved. Inspecting the Markov chain, we see that q' satisfies

$$q' = \frac{\mu}{\mu + \lambda} \left(\frac{1}{3} \right) + \frac{\lambda}{\mu + \lambda} (q')^2.$$

Let $\rho = \lambda/\mu$. Knowing $q' \in (0, 1)$, we can pick the correct solution to the quadratic,

$$q' = \frac{1}{2\rho} \left(1 + \rho - \sqrt{(1 + \rho)^2 - \frac{4}{3}\rho} \right).$$

As expected, this is decreasing in ρ : the more likely arrivals are compared to departures, the more likely it is that we transition from sad to happy at some point before going left.

We now compute the expected reward accumulated per renewal cycle, which we right as R . Let R_i be the total reward accumulated while going left from $i \geq 0$, and similarly for R'_i . This includes both the reward from state i or i' and reward from states to the right that are visited before going left. The reward rate for state i is $z^{(i-2)^+}$ (that is, $z^{\max\{i-2, 0\}}$), and the reward rate for state i' is $z^{(i-1)^+}$.

Examining the Markov chain, we see that

$$\begin{aligned} R &= \frac{1}{\lambda} + \frac{2}{3}R_1 + \frac{1}{3}R'_1 \\ R_1 &= \frac{1}{\mu + \lambda} + \frac{\lambda}{\mu + \lambda}(R_2 + R_1) \\ R'_1 &= \frac{1}{\mu + \lambda} + \frac{\lambda}{\mu + \lambda}(R'_2 + qR_1 + q'R'_1). \end{aligned}$$

Because the Markov chain repeats after 2 and 2', we know $R_3 = zR_2$ and $R'_3 = zR'_2$, giving us

$$\begin{aligned} R_2 &= \frac{1}{2\mu + \lambda} + \frac{\lambda}{2\mu + \lambda}(1 + z)R_2 \\ R'_2 &= \frac{z}{\mu + \lambda} + \frac{\lambda}{\mu + \lambda}((q' + z)R'_2 + (1 - q')R_2). \end{aligned}$$

This is the same q' as in the computation for L , so this is just a system of linear equations.

Sparing the reader the remaining details, we skip to the simple closed-form solution,

$$\hat{N}(z) = \frac{(\rho - 2)(-3\rho + X + 3)(\rho(z - 1) - 1)(\rho^2(X(z - 1) + 4z - 7) + \rho(X(z - 2) + 3z - 8) - 2(X + 3) + 3\rho^3(z - 1))}{(3\rho^2 + 5\rho + (\rho + 2)X + 6)(\rho z - 2)(X + \rho(3 - 6z) + 3)},$$

where $X = \sqrt{9\rho^2 + 6\rho + 9}$. We leave using this result to compute $E[N]$ and other metrics as a simple exercise for the reader.

3.3 Generalization to M/M/5/C2UPN

The same general approach works for longer rows of urinals, but the layer structure becomes more complex. Instead of solving a quadratic for q' , we must solve a system of quadratics to find several probabilities. Remarkably, the next-largest interesting case, the M/M/5/C2UPN, admits a closed-form solution. One probability is

$$\frac{1}{2\rho} \left(1 + \rho - \frac{1}{\sqrt{219}} \sqrt{111\rho^2 + 36\sqrt{9\rho^4 + 3\rho^3 + 3\rho + 9} + 55\rho + 111} \right),$$

and the rest are algebraic functions of it. From the probabilities, solving for L and R is routine. Whether the M/M/k/C2UPN has such an elegant solution for general k is a rich area for future research.

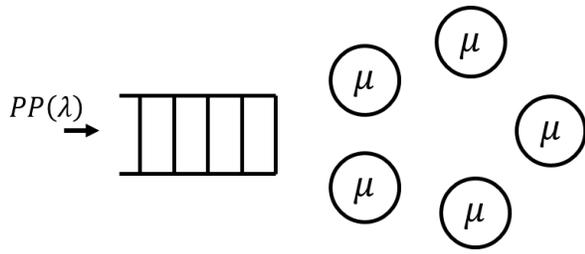


Figure 4.1: A 5-urinal CUP system.

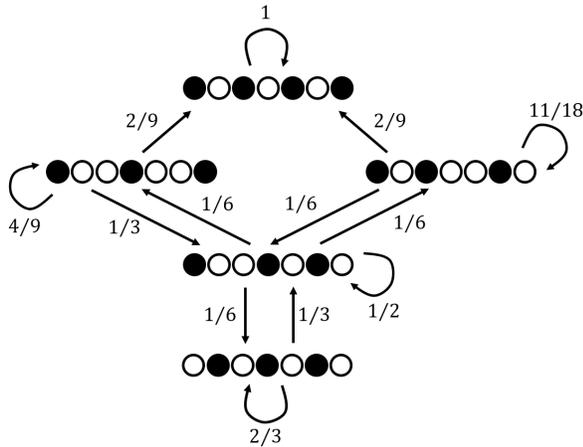


Figure 4.2: State transitions in the 7-state linear urinal configuration when the queue is non-empty.

4 ALTERNATIVE URINAL TOPOLOGIES

Thus far we have considered only the “standard” bathroom topology in which the urinals are arranged in a row, making the unease graph a path. But a multitude of urinal arrangements are possible. In this section we extend our results to an important class of alternative topologies: *Circular Urinal Positioning* (CUP).

In a CUP system, the urinals are arranged in a circle around a central pillar (see Figure 4.1), making the unease graph a cycle. Urinators arrive to the system as a Poisson process with rate λ and the urination duration is exponentially distributed with rate μ . Unlike in the line topology studied in the previous section, in a CUP system all of the urinals are symmetric in that there is not an endpoint with only a single neighboring urinal. While the CUP configuration makes it more challenging for urinators to find private urinals, we find that surprisingly, the CUP configuration makes analysis much more tractable in large bathrooms.

Consider the 7-urinal system. In a linear configuration, there are five possible states when the queue is non-empty and assuming perfect etiquette. The transitions between these states are shown in Figure 4.2. The complicated state transitions, and in particular the non-DAG structure, make it difficult to solve the resulting Markov chain using RRR. This is because it is possible to transition back and forth between pairs of states, meaning that finding the “leftward” probabilities will require solving a high-degree polynomial.

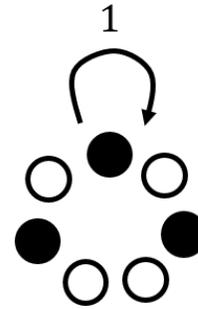


Figure 4.3: State transitions in the 7-urinal CUP configuration when the queue is non-empty.

Now suppose that instead our 7 urinals are arranged in a CUP structure. Now there is only one possible state when the queue is non-empty and all urinators observe perfect etiquette. The new state transition diagram is shown in Figure 4.3. Indeed, the system reduces to an M/M/3, which is easy to solve.

In general, the state space for any size urinal system is simpler in the CUP configuration than in the linear configuration. Due to the increased analytical tractability of large CUP systems, we suggest that all bathrooms be reconfigured so that the urinals satisfy a CUP structure.

5 CONCLUSIONS AND FUTURE WORK

In this paper we derived the first exact analysis of mean response time in urinal systems. Our approach, which uses the Recursive Renewal Reward technique, applies in the 3- and 5-urinal linear C2UPN systems, as well as in larger CUP systems.

There are several interesting and important directions for future work. Here we only consider the perfect etiquette setting, in which urinators never occupy adjacent urinals. When load is high, it may be necessary to move instead to the partial etiquette setting, in which urinators may occupy adjacent urinals if their urgency is sufficiently high. This complicates our Markov chain analysis because it introduces many new possible states for the urinal system. Furthermore, in the “probabilistic urgency” (p -urgency) setting, the exponential urination duration assumption may not be realistic. Empirical work has shown that a more realistic distribution is the sum of a delay before the start of urination, and a urination duration; both of these components depend on the proximity of the urinator to other urinators [7]. Hence in the partial etiquette/ p -urgency setting, we also need to extend our results to general urination duration distributions.

An alternative direction for future work involves the strategic decision of which urinal to choose in a p -urgency system. We have assumed that a urinator will choose uniformly at random from among the permissible urinals, but this need not be the case. A common strategy is to choose the urinal that maximizes the distance between urinators. While this strategy is beneficial for ensuring one’s individual privacy, it may reduce the overall system efficiency. An interesting direction for future work would be to investigate the Privacy-Efficiency Envelope.

We hope that this paper will serve as the start of a steady stream of future work on analyzing the performance of urinal and other crucial lavatory-related queueing systems.

REFERENCES

- [1] Michelle A Baillie, Shawndel Fraser, and Michael J Brown. 2009. Do women spend more time in the restroom than men? *Psychological reports* 105, 3 (2009), 789–790.
- [2] Spencer E Cahill, William Distler, Cynthia Lachowetz, Andrea Meaney, Robyn Tarallo, and Teena Willard. 1985. Meanwhile Backstage Public Bathrooms and the Interaction Order. *Journal of Contemporary Ethnography* 14, 1 (1985), 33–58.
- [3] Michelle Castillo. 2016. Study reveals what people do with their phones in the bathroom. *CNBC* (July 2016).
- [4] Anshul Gandhi, Sherwin Doroudi, Mor Harchol-Balter, and Alan Scheller-Wolf. 2013. Exact Analysis of the M/M/K/Setup Class of Markov Chains via Recursive Renewal Reward. In *Proceedings of the ACM SIGMETRICS/International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS '13)*. ACM, New York, NY, USA, 153–166. DOI: <http://dx.doi.org/10.1145/2465529.2465760>
- [5] Jeremy C Justus. 2006. Piss Stance: Private Parts in Public Places: An Analysis of the Men's Room and Gender Control. *Studies in Popular Culture* 28, 3 (2006), 59–70.
- [6] Evangelos Kranakis and Danny Krizanc. 2010. The urinal problem. In *International Conference on Fun with Algorithms*. Springer, 284–295.
- [7] R Dennis Middlemist, Eric S Knowles, and Charles F Matter. 1976. Personal space invasions in the lavatory: Suggestive evidence for arousal. *Journal of personality and social psychology* 33, 5 (1976), 541.
- [8] K. H. MOORE, D. H. RICHMOND, J. R. SUTHERST, A. H. IMRIE, and J. L. HUTTON. 1991. Crouching over the toilet seat: prevalence among British gynaecological outpatients and its effect upon micturition. *BJOG: An International Journal of Obstetrics and Gynaecology* 98, 6 (1991), 569–572. DOI: <http://dx.doi.org/10.1111/j.1471-0528.1991.tb10372.x>
- [9] Clint Rainey. 2015. Everything We Know About Human Bathroom Behavior. *NYMag* (May 2015).



CONFIDENTIAL COMMITTEE MATERIALS

SIGBOVIK 2017 Paper Review

Paper 75: RRR for UUU: Exact Analysis of Pee Queue Systems with Perfect Urinal Etiquette

I. P. Freeley, Low-brow humor expert

Rating: X

Confidence: Shaky

This very detailed and well-researched paper promises to raise the state of the art in SIGBOVIK bathroom humor. I was unable to evaluate the technical contribution because, after many years as a SIGBOVIK reviewer, I am unaccustomed to seeing real math. However, there were a number of avenues of research which were not discussed in the paper, which could make the submission more complete or at least be valuable considerations for future work. For example, the paper considers only the discrete case. I wonder how the paper's results would generalize to the continuous case:



The Next 700 Type Systems

Carlo Angiuli

March 31, 2017

Type systems classify programs in a way that enables compositional reasoning about their behavior. As a result, type systems have found a place as one of the major organizing principles of modern programming languages; much programming language research focuses on new ways of classifying programs in order to capture more sophisticated invariants, including dependent, gradual, refinement, linear, intersection, and existential types.

However, I feel that modern type systems focus too narrowly on *classification*, which is but one of the twenty-one definitions of the noun *type* in the Oxford English Dictionary. This myopic view of types has impeded the vast majority of the possible subdisciplines of type theory, as depicted in Figure 1.

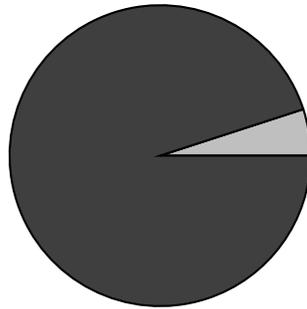


Figure 1: 95.2% of the definitions of *type*, *n.*, have not been explored in the context of type systems.

In the remainder of this paper, we describe a family of unimplemented type systems that is intended to span differences of meaning by a group of disparate frameworks (Landin, 1966).

Symbol, emblem. This sense of type theory is also known as *symbology*, a lesser-known field whose most famous researcher is Robert Langdon (Brown, 2000). Although Langdon has found success focusing his efforts toward the Illuminati and Catholic Church, it appears that other, less dangerous applications of symbology remain underexplored.

A pattern stamped onto the face of a coin. In this sense, type systems are methods of stamping patterns on coins. The earliest type systems required manually hammering a piece of metal between two dies. Improvements in metal-working and industrial technologies has enabled extensive automation in modern type systems. A recent theoretical advance in type systems occurred during the 2013 United States debt-ceiling crisis, in which some economists suggested that the United States Treasury could mint a \$1 trillion platinum coin in order to keep the government afloat without raising the debt ceiling (Matthews, 2013).

Tip. The theory of tips was famously studied by medieval theologians, who sought to compute the number of angels that could simultaneously dance on the head of a pin. (Although this has not been experimentally validated, Aquinas (1274) suggested that two angels cannot be in the same place.) While there are no known applications of tip theory, advances could lead to new transistor technologies. Unfortunately, funding sources are unlikely to surface.

A small block bearing a raised character, for use in printing. Modern type systems were invented by Johannes Gutenberg in 1440, and within decades, made an enormous impact on European society. I suggest that programming language researchers take credit for this early technological breakthrough in type systems.

The sort of person to whom one is attracted (*one's type*). It is well-known that types are (perhaps most) useful as tools for specifying interfaces at abstraction boundaries. Types also, apparently, apply at attraction boundaries; further research is warranted.

Printed characters (*in type*). \TeX is the most popular type system among traditional type theorists. When coupled with its large ecosystem of packages, \TeX 's type system is both expressive and aesthetically-pleasing, contrary to the popular belief that it has no types, and in fact lacks any facilities for abstraction.

An imperial edict released by Emperor Constans II in AD 648 prohibiting discussion of monothelism. The Type of Constans was a ban on the debate between monothelism and dyothelism—whether Jesus Christ, having both divine and human nature, possessed a single will, or two wills (divine and human).

In fact, traditional type systems are already ideal for restricting inquiry into the nature of things; recall, for instance, the fable of Professors Descartes and Bessel in Reynolds (1983). To a programming language researcher, it is clear that the Type of Constans is parametrically polymorphic in the will of Christ, and the prohibition of discussion is simply a free theorem (Wadler, 1989).

The specimen originally used to name a species (*type specimen*). This is clearly just a mode of use of singleton types; given a specimen s , its species is the denotation of the type $S(s)$ of all specimens equal to s .

type, v. To write with a keyboard. Although debate rages on over which type system is best—QWERTY, Dvorak, Colemak, et cetera—most evidence remains anecdotal. This debate has not impacted the success of QWERTY and its close relatives (Noyes, 1983), but the field is nevertheless in dire need of rigorous theoretical study. Given the relevance of keyboards to both proving and programming, I suggest submitting research on this subject to the annual TYPES International Conference on Types for Proofs and Programs.

References

- [1] Thomas Aquinas. *Summa Theologiæ*. 1274.
- [2] Dan Brown. *Angels & Demons*. New York: Simon and Schuster, 2000. ISBN: 978-0-7434-1239-1.
- [3] P. J. Landin. “The Next 700 Programming Languages”. In: *Commun. ACM* 9.3 (Mar. 1966), pp. 157–166. ISSN: 0001-0782. DOI: 10.1145/365230.365257. URL: <http://doi.acm.org/10.1145/365230.365257>.
- [4] Dylan Matthews. “Michael Castle: Unsuspecting godfather of the \$1 trillion coin solution”. In: *The Washington Post* (Jan. 2013). URL: <https://www.washingtonpost.com/news/wonk/wp/2013/01/04/michael-castle-unsuspecting-godfather-of-the-1-trillion-coin-solution>.
- [5] Jan Noyes. “The QWERTY keyboard: a review”. In: *International Journal of Man-Machine Studies* 18.3 (1983), pp. 265–281. ISSN: 0020-7373. DOI: [http://dx.doi.org/10.1016/S0020-7373\(83\)80010-8](http://dx.doi.org/10.1016/S0020-7373(83)80010-8). URL: <http://www.sciencedirect.com/science/article/pii/S0020737383800108>.
- [6] John C. Reynolds. “Types, abstraction, and parametric polymorphism”. In: *Information Processing* (1983), pp. 513–523.
- [7] Philip Wadler. “Theorems for Free!” In: *Proceedings of the Fourth International Conference on Functional Programming Languages and Computer Architecture*. FPCA ’89. Imperial College, London, United Kingdom: ACM, 1989, pp. 347–359. ISBN: 0-89791-328-0. DOI: 10.1145/99370.99404. URL: <http://doi.acm.org/10.1145/99370.99404>.

A Modular Approach to State-of-the-Art Big Data Visualization

Keith A. Maki

Abstract

In our technical demo, we will present a novel approach to visibly large data visualization using a modular segmented display design. We demonstrate that this approach is both large and displays data, and we compare to more established methods of big data display, finding that the demonstrated approach can display very big data indeed.

Introduction

Recently, there has been a stir in the computational research community over the need for research conducted on big data. This is an incredibly promising area, and much work can surely be done provided the data is big enough. However, most computer scientists work at desks with a small number of monitors that are only so big, which severely limits the size of the data they can work with. We believe that to empower the field of big data science new technologies must be developed to meet the demand for ever-increasing sizes of data. Our technical demonstration will provide a novel application of common indicator technologies in a modular configuration highly amenable to scalable visualization of really big data. We compare the presented design to several alternative approaches of varying degrees of conventionality and we demonstrate its ability to visualize both really big numbers and really big text in a variety of colors.

Background and Related Work

The needs of the big data research community have rarely been the focus of computer engineers developing computing technologies, which have long been optimized for computational capabilities, power efficiency, portability, and cost. Ironically, as electronics capabilities have improved, data storage has become more and more limiting from a big data standpoint, such that in order to store data on a standard 3.5in hard drive, data is physically constrained to have a surface area no larger than the largest planar cross section of the case volume, which is not very large. Gone are the days when magnetic tape reels the size of hubcaps stored data, although to be fair in order to store data larger than the tape width a method such as striping had to be used, and putting the data back together for any practical purposes had in our experience always resulted in a tangled mess.

To overcome these limitations, data is generally analyzed with the help of a visualizer, such as a computer monitor, that attempts to compensate for the smallness of the data by biggifying it so it can be worked with as though it were big data. However, these technologies are limited in their capabilities and can only compensate so much, leading to problematic results such as that shown in Figure 3.

Nevertheless, a number of technologies have attempted to provide solutions to these challenges. For example, the electronic scoreboard is perhaps the most familiar and well-known example of a big data visualizer. These displays are typically implemented using incandescent or LED technology in either segmented, eggcrate, or dot matrix format, although flip-disc and vane display technologies have also been used. These displays can display very big data but are often bulky and impractical for transporting, and are capable of visualizing only a fixed number, format, and color of numeric digits, which limits their usefulness in many big data applications, particularly text.

Similarly, projection systems have also found their way into standard use, and these overcome the limitations of scoreboards on data format, color, and even size, being able to project colored numeric and text data at sizes proportional to the distance between the projector and the visualization surface. However, the data visualizations produced by these systems suffer from a noticeable dimming in brightness inversely proportional to their area. Additionally, while very large, crisp projection systems, such as IMAX¹, exist which are capable of visualizing big data with thousands of digits, the scalability in terms of the number of digits is limited by the physical dimensions of the projection surface and the optical resolution properties of the projector, and there will be a point at which the data is either too small or too indistinct to reliably work with. Additionally, such research-grade big data projection systems are generally not portable, which further detracts from their utility.

The presented system will attempt to address all of these shortcomings, providing a portable, scalable design using modular display panels. Our work is most similar to that of Overpass Light Brigade², who make use of LED signboards to display the individual characters that make up a piece

¹<http://www.imax.com/>

²<http://overpasslightbrigade.org/>

of data. However, where OLB uses monochrome LEDs in hardwired letter configurations, our modular display panels are both programmable and wicked cool, allowing for dynamic visualization of both numeric and text data in variable color.

Proposed Design

We propose the use of LED indicators positioned in a segmented display pattern of the desired size. The LED layout for a single panel is shown in Figure 1. This configuration was chosen to allow for compatibility with standard 16-segment display layout alphanumeric characters, but the design is flexible enough to allow for similar but modified layouts. For our prototype system, we make use of WS2812B RGB LEDs, which have an integrated shift register to manage PWM control for each pixel separately. We wire the LEDs such that each panel has the same configuration, with a data in port on one side and a data out port on the other. In this way, color programming information can be shifted in down a chain of several panels. For the LED server node, we use a BeagleBoneBlack running the LEDscape³ framework, to handle the buffering of several chains of LEDs simultaneously; and we send frames to the server over an ethernet connection using the Open Pixel Control⁴ communications protocol.

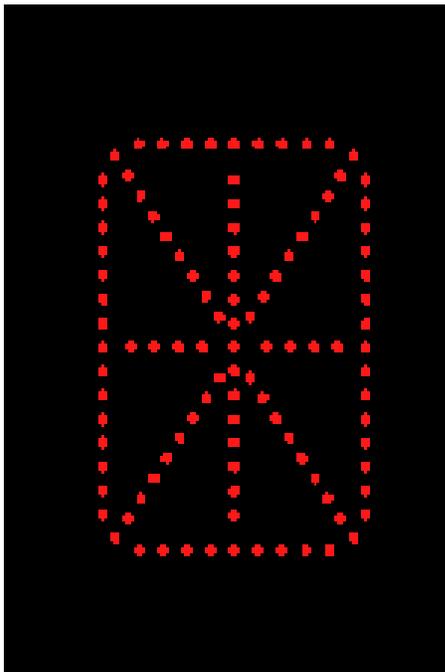


Figure 1: The layout of LEDs on each panel

³<https://github.com/Yona-Appletree/LEDscape/>

⁴<http://openpixelcontrol.org/>



Figure 2: Big data scientist conducting a visualization test

Experiments

We compared the proposed approach with a number of alternative technical solutions for big data visualization. Each approach was coded by two raters for four binary attributes:

1. **Big data** Whether the technology constitutes big data visualization.
2. **Portable** Whether it's feasible to transport the technology on a semi-regular basis.
3. **Scalable** Whether the technology scales for big data with an arbitrary number of characters.
4. **Nerdy factor** Whether people will know you're a real nerd if you say you visualize things with the technology.

The inter-rater agreement was strong (Cohen's $\kappa = 0.85$). The results of our comparisons are shown in Table 1.

We have manufactured eight prototype panels, controlled using four panels on each of two channels on one node to demonstrate the potential of our system, but the prototype setup is scalable to five panels on each of 48 channels on each of at least 16381 nodes, for a total of altogether too many panels⁵. With each digit measuring at nearly two feet tall, our prototype panels definitely constitute "big", and we've had a lot of fun using them to visualize data in several colors and combinations. An image of a big data scientist conducting a data visualization test using the prototype system is shown in Figure 2.

Conclusion and Future Work

In this report, we have described the capabilities of our novel big data visualization technology. We have explored the technical advantages of our system when compared to other data visualization technologies with respect to state-of-the-art big data research desiderata. Future work will explore more freely configurable pixel layouts such as modular

⁵We did the math.

Technology	Big data	Portable	Scalable	Nerdy factor	Notes
Beach	Yes	No	Yes	No	Not programmable
Calculator	No	Yes	No	Yes	Not a big data solution
IMAX	Yes	No	No	No	I doubt your lab has space for one
Projector	Yes	Yes	No	Yes	Brightness varies with size of data
Scoreboard	Yes	No	No	Yes	Limited data format
Ulexite	No	Yes	No	Yes	Only a rocky solution
Our system	Yes	Yes	Yes	Yes	Fails insignificance test ($p = i$)

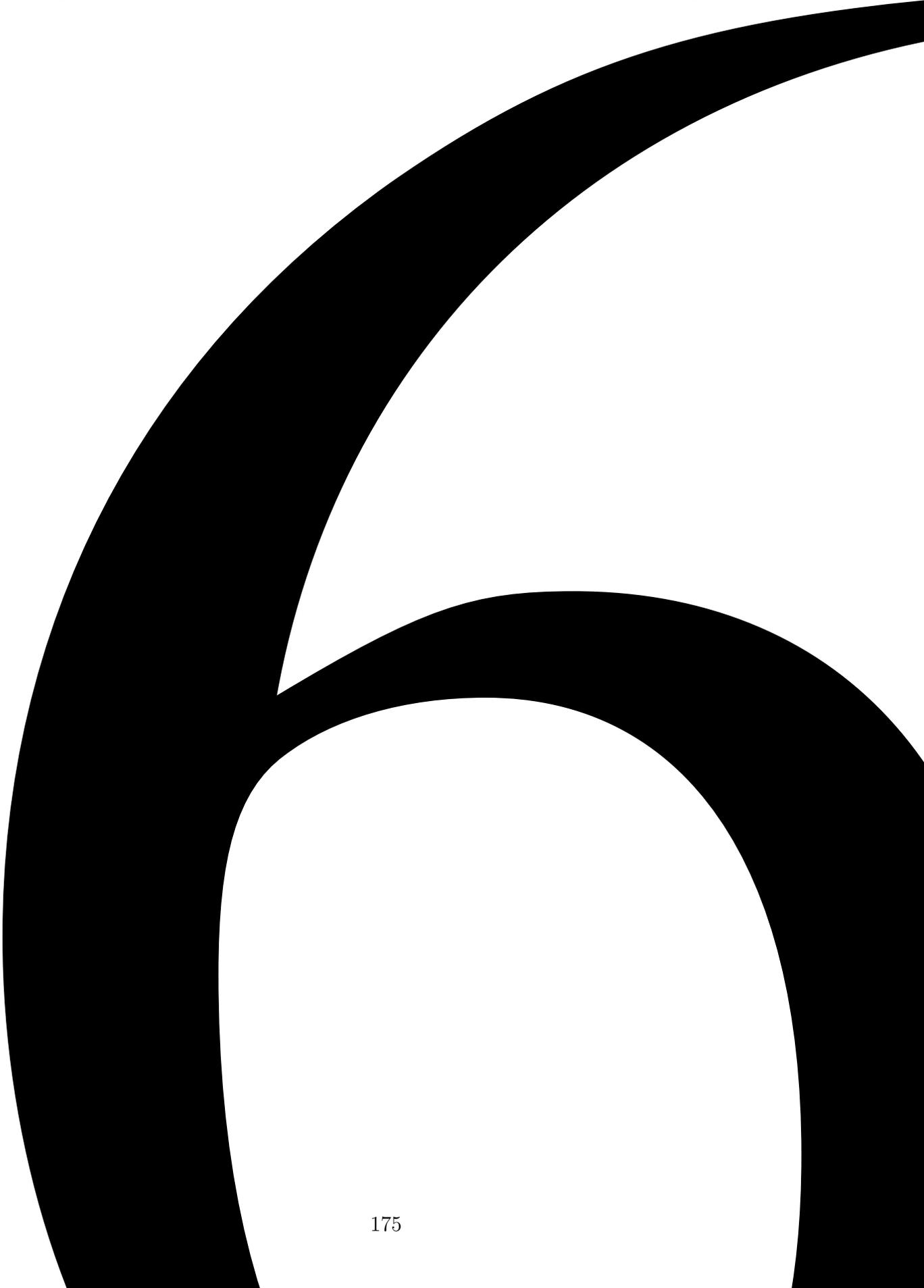
Table 1: A comparison of the proposed approach with various data visualization technologies

dot matrix designs, allowing for flexible data size stretching multiple panels in any direction. Additionally, it would be interesting to explore the use of drone swarm technology to visualize big data in higher dimensions, but we definitely would need grant money for that.

Acknowledgements

Special thanks to Bob Rudenborg and Sam Selders for all their help with the soldering.

Figure 3: A piece of data which likely exceeds the physical limits of your display technology



Efficient Computation of an Optimal Portmantout

David Renshaw

31 March 2017

1 Introduction

A *portmantout* is a string composed of sequentially overlapping words from a word set \mathbf{L} , such that each word from \mathbf{L} appears in the string at least once. For example, if $\mathbf{L} = \{\text{an, no, on}\}$, then `anon` is a portmantout for \mathbf{L} ; in fact it is the *only* portmantout for \mathbf{L} . For larger \mathbf{L} , there are often many portmantouts of differing lengths, suggesting a natural question: how short of a portmantout can we construct? Traditionally, interest in this question has centered on a particular 108,709-word word set called `wordlist.asc` [3]. The first published portmantout for `wordlist.asc` had length 630,408 [5]. Later, new methods were used to find a portmantout for `wordlist.asc` of length 537,136 and to prove any portmantout for `wordlist.asc` must have length at least 520,732 [6].

In this paper we present an efficient method for finding an *optimal* portmantout. Although the method is not guaranteed to work on all word sets, it does succeed on `wordlist.asc`, finding an optimally short portmantout of length 536,186 using less than fifteen seconds of computation time. The method is due to Anders Kaseorg, who developed it while solving a related programming puzzle [1] [2].

2 The Method

We encode the portmantout problem as a *minimum-cost network flow* problem. In general, a network flow problem has the following input data:

- A set N of nodes.
- A set A of directed arcs between nodes. Each arc $a \in A$ has an input node $a_{\text{in}} \in N$ and an output node $a_{\text{out}} \in N$.
- A supply map: each node $n \in N$ has associated an integer b_n , representing the “supply” of that node. If b_n is positive, then n is a “source” node. If b_n is negative, then n is a “sink” node.
- A cost map: each arc $a \in A$ has associated an integer c_a representing the cost per unit of flow along that arc.
- A capacity map: each arc a has a capacity $u_a \geq 0$ representing the maximum flow allowed through that arc.

Given this data, we seek a flow x on A to minimize the cost $\sum_{a \in A} c_a x_a$, subject to:

$$\sum_{a \in A, a_{\text{in}}=n} x_a - \sum_{a \in A, a_{\text{out}}=n} x_a = b_n$$
$$0 \leq x_a \leq u_a$$

As is expounded in [4], there are efficient algorithms for solving such problems. Moreover, if the supply, cost, and capacity values are all integers, then there is guaranteed to be an optimal solution vector x that is also entirely integral.

2.1 Encoding the Portmantout Problem

Given a word set \mathbf{L} , we construct a minimum-cost network flow problem as follows:

- For each word $w \in \mathbf{L}$, we create a “start” node $\llbracket w \rrbracket$ and an “end” node $\lvert w \rvert$. If the word is redundant (i.e. is contained in some other word in \mathbf{L}) then these have zero supply. If the word is not redundant, then $\llbracket w \rrbracket$ has supply $b_{\llbracket w \rrbracket} = -1$ and $\lvert w \rvert$ has supply $b_{\lvert w \rvert} = 1$.

- For each $w \in \mathbf{L}$, we create an arc with zero cost from $\llbracket w \rrbracket$ to $|w\rangle$. This is called the “extra” arc. It allows us to reuse words more than once as connectors between other words.
- For each string s that is either a prefix or a suffix of any word in \mathbf{L} , we create an “affix” node \bar{s} . These nodes represent the overlap between successive words in a portmantout.
- For each $w \in \mathbf{L}$ and each affix s , if s is a prefix of w then we create an arc from \bar{s} to $\llbracket w \rrbracket$ with cost equal to the length of w minus the length of s . The cost represents the number of letters contributed to the length of the final portmantout.
- For each $w \in \mathbf{L}$ and each affix s , if s is a suffix of w then we create a zero-cost arc from $|w\rangle$ to \bar{s} .
- We create an “initial” node $|\emptyset\rangle$ with supply $b_{|\emptyset\rangle} = 1$, and for each $w \in \mathbf{L}$ an arc from $|\emptyset\rangle$ to $\llbracket w \rrbracket$ with cost equal to the length of w .
- We create one “final” node $\llbracket \emptyset \rrbracket$ with supply $b_{\llbracket \emptyset \rrbracket} = -1$, and for each $w \in \mathbf{L}$ an arc from $|w\rangle$ to $\llbracket \emptyset \rrbracket$, with cost 0.
- We set the capacity of every arc to ∞ .

2.2 Recovering a Portmantout

We plug the encoded problem into a solver and get back an optimal integral flow x . Our objective now is to read off a portmantout from x by following the flow from the initial node $|\emptyset\rangle$ to the final node $\llbracket \emptyset \rrbracket$, allowing a jump from $\llbracket w \rrbracket$ to $|w\rangle$ for each non-redundant w . To formalize this objective, we create a new directed graph H with the same nodes N as before. For each unit of flow along an arc in x , we draw a distinct arc in H between the same nodes. For example, if $x_a = 3$ then we draw three distinct arcs in H from a_{in} to a_{out} . In addition, for each non-redundant $w \in \mathbf{L}$, we draw an arc from $\llbracket w \rrbracket$ and $|w\rangle$, i.e. from that word’s sink node to its source node. Finally, we draw an arc from $\llbracket \emptyset \rrbracket$ to $|\emptyset\rangle$. We then try to find an Eulerian circuit on H . That is, we look for a path that visits each arc exactly once and ends up where it started. If we can find such a circuit, then we have found an optimal portmantout and we are done.

Unfortunately, H is not guaranteed to have a Eulerian circuit. Consider, for example, the case where $L = \{\text{abyz}, \text{yzab}, \text{zxy}\}$. Then zxyzabyz , which has length 8, is the shortest possible portmantout for L , but the minimum-cost flow algorithm finds a flow of cost 7 with abyz and yzab in a cycle, producing a graph that has no Eulerian circuit. In such cases, the flow x gives a lower bound on the length of any portmantout for \mathbf{L} .

3 Future Work

Is the problem of searching for an optimally-short portmantout NP-hard in general? A plausible-looking equivalence with the Traveling Salesman Problem was presented in [5], but its reduction from TSP uses a unary encoding of the costs of edges, and therefore can cause the size of problems to grow exponentially, which apparently invalidates the proof. Can that proof be repaired? Or, perhaps, can the algorithm described in the present paper be amended to handle all cases?

References

- [1] Compounding english. <https://codegolf.stackexchange.com/questions/87311/compounding-english/87534>.
- [2] shortmantoutmost. <https://github.com/andersk/shortmantoutmost>.
- [3] wordlist.asc. <http://www.cs.cmu.edu/~tom7/portmantout/wordlist.zip>.
- [4] Ravindra K. Ahuja James B. Orlin and Thomas L. Magnanti. *Network Flows: Theory, Algorithms, and Applications*. Pearson, 1993.
- [5] Tom Murphy VII Ph.D. The portmantout. In *Proceedings of SIGBOVIK 2015*, 2015.
- [6] David Renshaw and Jim McCann. A shortmantout. In *Proceedings of SIGBOVIK 2016*, 2016.

A Appendix: An Optimal Portmantout for wordlist.asc



Raccoon track

Talkin' Trash



- 32** Garbage collection for heaps only a mother could love
Your Parents
- 33** A new paradigm for robotic dust collection: Theorems, user studies, and a field study
Rachel Holladay and Siddhartha S. Srinivasa

Garbage Collection for Heaps Only a Mother Could Love

Your Parents*

March 31, 2017

Abstract

What do I write here? Is it just a summary of the paper? Isn't that what the introduction is for? I don't usually write papers in computer science conferences, I just know you're busy and thought this might be the best way to talk to you without distracting you from your work.

1 Introduction

Hi, it's me. I was looking through the basement trying to make some space. It's a real heap down there. Anyway, I came across some of your old stuff and I wasn't sure if you still needed it. Your father just wanted to throw it all away, but I thought I'd call just to check. This is a good time, isn't it? Are you sure? You'd tell me if it wasn't, right? OK, well, here goes.

2 Background

Where are you? I'm having a little trouble hearing you, there's a lot of noise in the background. Are you out walking? Oh, is it very windy there? I can call back if this isn't a good time. OK, that's better.

3 Algorithm

So, the first box is some of your old school things. Oh, it's your notes from AP Calculus. Remember that class? Your teacher, what was her name? Oh, come on, she was really nice. She called home that time you asked a question in class and she didn't know the answer? I know you remember. Anyway, I have your old notebook here. You're in school for computer science, you must need math a lot, right? Are you sure you won't need it? If you need it now, I can mail it. Well, if not, it'll be waiting for you here next time you come home. Oh, but then you might also need your review book from the AP exam. I think I saw a box of those old review books somewhere around here. Oh, and your pre-calc notebook is here too. You'll need that to make sense of the calc notes, right? I guess I should save that also. We need a system for this. I think I have an idea. See Figure 1.

4 Results

Well, there's a lot here. You know what? It all looks like it's stuff you might want to keep around. Well, you never know. If I move it all into your room, maybe next time you're home you can decide what you want to keep there and what you want to move into the attic.

5 Discussion

Enough about us. How are you? How's school, are you almost done? Why does it take so long for you to finish your PhD? You work with computers, shouldn't it be really fast? Are you eating alright? If you need more money, we can send you some. I just want to make sure you're eating well and everything is OK. Is everything OK?

*Current contact information: You know full well what it is, you just don't use it often enough

```

proc mark(roots)
  start = time.now()

  while not (roots.empty)
    if (time.now() - start > time_limit)
      print "I give up, can you take a look at this stuff?"
      break

    loc = roots.removeOne()
    loc.mark()
    for (loc ' in loc.pointers())
      roots.add(loc ')

```

Figure 1: We can mark all of the things that you need, and then I can come back later to sweep up.

6 Conclusion

Well, I imagine you need to get back to work. Can we talk again this weekend?

A Future Work¹

While this discussion was (mercifully) short, it leaves open many rich areas of potential future research. In the future, your parents hope to investigate many other collection strategies from the literature, such as parallel (“your father and I both spent the whole day on this”), concurrent (“don’t mind me, keep doing what you’re doing while I clean up your mess”), incremental (“just clean up after yourself as you work, why is that so hard?”), real-time (“while you’re home, you can take care of this now”), tagless (“are these boxers yours or your father’s?”) and copying (“Mom, she’s copying me!”). Many cultures also provide for unique collection strategies worthy of discussion. For example, this paper does not discuss the frum-space invariant of Jewish garbage collectors (“What’s this ramen doing here, don’t you know it’s Pesach?”). I’d write more, but oops, my dad is calling to bug me about my car insurance.

¹Appendix by Stefan Muller, School of Computer Science, Carnegie Mellon University

A New Paradigm for Robotic Dust Collection: Theorems, User Studies, and a Field Study

Rachel Holladay
Robotics Institute
Carnegie Mellon University
rmh@andrew.cmu.edu

Siddhartha S. Srinivasa
Robotics Institute
Carnegie Mellon University
siddh@cs.cmu.edu

Abstract—We pioneer a new future in robotic dust collection by introducing passive dust-collecting robots that, unlike their predecessors, do not require locomotion to collect dust. While previous research has exclusively focused on active dust-collecting robots, we show that these robots fail with respect to practical and theoretical aspects, as well as human factors. By contrast, passive robots, through their unconstrained versatility, shine brilliantly in all three metrics. We present a mathematical formalism of both paradigms followed by a user study and field study.

I. INTRODUCTION

There has been renewed recent interest in the design of efficient and robust dust-collecting robots [8, 4]. The oppression of constant dust raining over our heads calls out for immediate attention. Furthermore, the increased cost of legal human labor, and increased penalties for employing illegal immigrants, has made dust-collection all the more critical to automate [7, 10].

However, all of the robotic solutions have focussed exclusively on what we define (see Definition 2 for a precise mathematical definition) as *active dust-collecting robots*. Informally, these are traditional robotic solutions, where the robot locomotes to collect dust. It is understandable why this seems like a natural choice as humans equipped with vacuum cleaners are, after all, also active dust-collectors.

Unfortunately, active dust collection presents several challenges: (1) *Practical*: they require locomotion, which requires motors and wheels, which are expensive and subject to much wear and tear, (2) *Theoretical*: most active dust-collectors are wheeled robots, which are subject to nonholonomic constraints on motion, demanding complex nonlinear control even for seemingly simple motions like moving sideways [3, 2, 11], (3) *Human factors*: several of our users in our user study expressed disgust, skepticism, and sometimes terror, about the prospect of sentient robots wandering around their homes, for example:

I don't want a f*cking robot running around all day in my house.

	Active	Passive
Practical		
Theoretical		
Human Factors		

TABLE I: Passive dust-collection outperforms active dust-collection in all metrics

In this paper, we propose a completely new paradigm for dust collection: *passive dust-collecting robots* (see Definition 1 for a precise mathematical definition). Informally, these are revolutionary new solutions that are able to collect dust *without any locomotion!*

As a consequence, passive dust-collecting robots address all of the above challenges: (1) *Practical*: Because they have no moving parts like wheels or motors, they are both inexpensive and incur no wear and tear, (2) *Theoretical*: because passive dust-collectors can be trivially parallel transported to the identity element of the $SE(2)$ Lie Group, they require no explicit motion planning (in situations where parallel transport is inefficient, the robot can be physically transported to the identity element), (3) *Human Factors*: as passive dust-collecting robots are identical to other passive elements in our homes and work places (like walls, tables, desks, lamps, carpets), their adoption into our livespace is seamless.

In addition, we present and analyze a mathematical model of dust collection. Using our model, we can, for the first time, answer which robot-type is more efficient. This is a critical question to consider in order to inform future cleaning choices.

Our analysis reveals that for a certain choice of constants, a passive dust cleaning robot is more efficient than its active counterpart. Through a user study, we contrast this with user's perceived perception of robot efficiency and what factors influence their choices.

To explore what choices are actually made we leveraged a field study of Carnegie Mellon's Robotics Institute

to determine the prevalence of each robot type. This study reveals that passive dust collecting allows for a wider range of morphologies, suggesting that passive dust collecting is a more inclusive characterization. Furthermore, we see that rather than two paradigms there is a continuum of dust collecting robots.

Our work makes the following contributions:

Mathematical Formulation. We present a model of active and passive dust collecting robots followed by an efficiency tradeoff analysis.

Preference User Study. We surveyed college students to determine what kind of robot they preferred and which they perceived to be more efficient.

Field Study. Using data on the robots of the Robotics Institute we investigate the more popular robotic paradigms.

We believe our work takes a first step in launching a new discussion concerning the nature of robotic dust collection, paving the way for future cleanliness.

II. A MATHEMATICAL MODEL FOR DUST COLLECTION

In order to compare and analyze active and passive dust collecting robots we present a mathematical model of their dust collection capabilities. With this model, we dare to ask: which robot is more efficient?

A. Dust Model

We model dust as a pressureless perfect fluid, which has a positive mass density but vanishing pressure. Under this assumption, we can model the interaction of dust particles by solving the Einstein field equation, whose stress-energy tensor can be written in this simple and elegant form

$$T^{\mu\nu} = \rho U^\mu U^\nu \quad (1)$$

where the world lines of the dust particles are the integral curves of the four-velocity U^μ , and the matter density is given by the scalar function ρ .

Remarkably, unless subjected to cosmological radiation of a nearby black hole, or a near-relativistic photonic Mach cone, this equation can be solved analytically, resulting in dust falling at a constant rate of α .

We model our robots as covering 1 unit² area of space-time. We present our models for passive and active robots before performing comparative analysis.

B. Passive Robot Model

We provide the following formalism:

Definition 1. We define a passive dust collecting robot as a robot that does not move, collecting the dust that falls upon it.

Lemma 1. The dust-collecting capability of a passive dust-collecting robot is given by

$$D_{\text{passive}} := \alpha \quad (2)$$

C. Active Robot Model

We provide the following formalism:

Definition 2. We define an active dust collecting robot as a robot that moves around the space, actively collecting dust.

We model our active robot as driving at speed β . We assume that our robot can only active collect dust of height h . This assumption is drawn from iRobot's Roomba, which reportedly can get stuck on cords and cables. As a simplifying assumption we will assume that the robot always collects dust of height h , implying that there is always at least dust of height h prior to the robot's operation.

Lemma 2. The dust-collecting capability of an active dust-collecting robot is given by

$$D_{\text{active}} := h\beta^3 + \frac{\alpha}{\beta} \quad (3)$$

Proof: It is obvious that the robot actively collects $h\beta^3$ dust.

However this is not the entire story. As the robot drives, actively collecting dust, it also passively collects the dust that happens to fall on it. To model this, we consider the robot passing over some fixed line. Some portion of the robot is occluding this line for $\frac{1}{\beta}$ seconds. Thus the robot passively collects $\frac{\alpha}{\beta}$ dust.

Thus, combining the active and passive components our active robot collects:

$$D_{\text{active}} := h\beta^3 + \frac{\alpha}{\beta} \quad \blacksquare$$

D. Model Comparison

We next compare for what tradeoffs there are between passive and active dust cleaning robots. We pose this as the question: When are passive dust cleaning robots more efficient than their active counterparts? Hence when is $D_{\text{passive}} > D_{\text{active}}$?

We are now ready to prove our main theorem.

Theorem 1. The dust-collecting capability of a passive robot exceeds the dust-collecting capability of an active robot when

$$\alpha > \frac{h\beta^4}{\beta - 1} \quad (4)$$

Proof: Using (2) and (II-C) we get:

$$D_{\text{passive}} > D_{\text{active}} \\ \alpha > h\beta^3 + \frac{\alpha}{\beta}$$

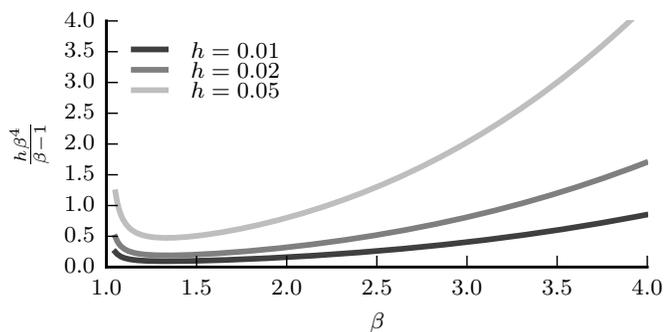


Fig. 1: Comparing active to passive collecting robots (fix caption)

With some simple arithmetic this becomes:

$$\alpha > \frac{h\beta^4}{\beta - 1}$$

Fig.1 shows this function over a variety of β s and a few choices of h . The y-axis can be viewed as a measure of efficiency. A passive robot's efficiency corresponds to a straight line across the y-axis at its α value.

As the h value increases, the active robot's efficiency increases, which follows from the fact that as it drives, it can collect more dust. While we see an initial drop in efficiency due to a β increase, owing to the fact that the active robot collects less dust passively, this effect is then dwarfed by a faster moving robot that can cover more ground.

III. USER STUDY

Having a developed a model of passive and active dust collecting robots we used a user study to evaluate people's opinions on each type of robots efficiency. This is critical in developing effective robots as we need to explore the possible discrepancies between perceived versus actual robot capability [1].

A. Experimental Setup

We created an online form to evaluate users opinions of passive and active dust collecting robots. Provided users with Definition 1 and Definition 2, we then asked them the following questions:

Question 1. Which type of robot do you think collects more dust: an active dust collecting robot or a passive dust collecting robot? Why?

Question 2. Which robot would you prefer to have?

For Question 2 the options were: Active dust collecting robot, Passive dust collecting robot, Whichever robot is the most efficient at collecting dust. Our goal in asking this was to determine what people value more, the illusion of efficiency or actual efficiency.

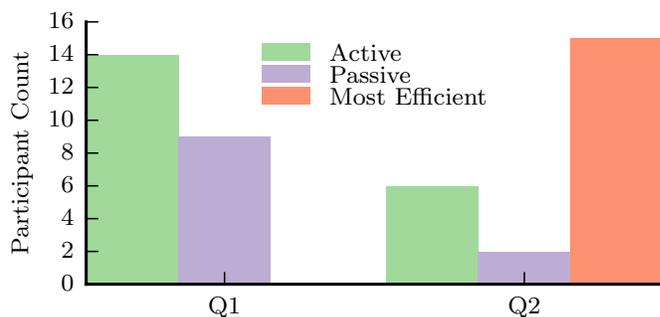


Fig. 2: User Study Results

Participants We recruited 23 Carnegie Mellon students (14 males, 9 females, aged 21-23) through online sources.

B. Analysis

The results of our user study can be seen in Fig.2. While people believe that the active robot collects more dust, people would prefer to have the most efficient robot, regardless of its capabilities.

What is perhaps more telling is the variety of user responses we had to why they believed each robot would collect more dust.

Those who supported passive dust collecting robots listed a variety of reasons, with many people concerns with active dusting robots dispersing and upsetting more dust than the collect. One user rationalized his choice by the nature of dust saying "I've observed that the stuff that collects the most dust in my place are the items that are static, therefore I would assume that the static robot might collect more dust."

Still other users took a more global view with one user, as mentioned above, claiming that they "don't want a f*cking robot running around all day" and another, accepting the harsh realities of time remarked "All robots ultimately become a passive dust-collecting robot."

For every supporter of passive robots, there were still more who argued for active robots. Almost every person, in explaining their choice, argued that active robots, due to their mobility, would be able to cover a larger space. This highlights the dichotomy between *efficiency and coverage*.

While our passive dust collecting robot can provide superior efficiency, its lack of locomotion greatly reduces is potential coverage. By contrast, the active robot has the ability to move around, coverage potentially all of the room, given some amount of time.

IV. FIELD STUDY

Given the results of our user study in Sec. III, we next probe into how these preferences are reflected in reality. Carnegie Mellon's Robotics Institute is home is a large variety of robots and using the 2010 robot census we



Fig. 3: Throughout the robots that can be seen at Carnegie Mellon’s Robotics Institute we see a variety of passive dust collecting robots that range widely in shape, size, initial function and even cost.

analyzed what kind of dust collecting robot we actually see [9].

Of the 261 robots listed on the census¹ with complete information, we see that none of them are designed to collect dust actively. However, we can assume many of them collect dust passively. Twenty were listed as having no mobility, making them official passive dust collecting robots. Even the eighty-six robots that have wheeled mobility are unlikely to be driving most of the time and therefore spend much of their life as passive dust collecting robots.

In fact, broadening out, despite the variety in morphologies and mobilities from wheeled to winged, from manipulation to entertainment to competition, most, if not all, of the robots at the Robotics Institute spend large quantities of their tenure as passive dust collecting robots. While active dust collecting robots are constrained by their function to have certain properties, passive dust collecting robotics is an all-inclusive, all-accepting genre that allows for nearly any characterization. We see a huge variety of robots in Fig.3. They can be old or new, outrageously expensive or dirt cheap, beautifully crafted or hastily thrown together.

Yet, if they do nothing, they all have the ability inside of them to be passive dust collecting robots. Given the guidelines provided by our model in Sec. II, these robots have the capacity to be more efficient than their try-hard active collection counterparts. Based on the results of our study (Sec. III) this makes them more desirable.

From these insights, it is now clear why the CMU Robotics Institute does not have any active dust collecting robots on record. They have been surpassed by their more efficient, more inclusive, more desirable counterparts: passive dust collecting robots.

¹The original census data was provided directly from its author.

V. DISCUSSION

While our analysis presented in Sec. II outlines two classes of robots, our field study from Sec. IV reveals a continuum of dust collecting robots. Robots that do not actively collect dust but are not entirely stationary, such as robots that are simply underused, represent the middle ground of dust collection. We can even think of air filters as dust collecting robots that actively collect dust but do not do so by moving themselves. This adds a new dimension of what it means for a robot to be active.

This work also aims to highlight the underappreciated advantages of passive dust collecting robots. Passive robots, unconstrained by a need for explicit dust collecting capabilities, afford a wide range of morphologies. This allows for incredibly flexibility in designing the possible human-robot interaction schemes, which is critical to a cleaning robot’s acceptance [6, 5].

While we focused on dust collecting robots are model generalizes to other situations, such as moving in the rain. Specifically, our model can be used to model whether you would get more wet by standing still or running through the rain.

We hope that this work will raise awareness for passive dust collecting robots and raise further discussion on the nature of dust collection.

ACKNOWLEDGMENTS

This material is based upon work supported by the infinite discretionary money-bag. We do not thank the members of the Personal Robotics Lab for helpful discussion and advice as this project was kept entirely super secret from them.

REFERENCES

- [1] Elizabeth Cha, Anca D Dragan, and Siddhartha S Srinivasa. Perceived robot capability. In *RO-MAN*, pages 541–548. IEEE, 2015.
- [2] Howie Choset. Coverage for robotics—a survey of recent results. *Annals of mathematics and artificial intelligence*, 31(1):113–126, 2001.
- [3] Nakju Lett Doh, Chanki Kim, and Wan Kyun Chung. A practical path planner for the robotic vacuum cleaner in rectilinear environments. *Transactions on Consumer Electronics*, 53(2), 2007.
- [4] Paolo Fiorini and Erwin Prassler. Cleaning and household robots: A technology survey. *Autonomous Robots*, 9(3):227–235, 2000.
- [5] Jodi Forlizzi and Carl DiSalvo. Service robots in the domestic environment: a study of the roomba vacuum in the home. In *SIGCHI/SIGART*, pages 258–265. ACM, 2006.
- [6] Bram Hendriks, Bernt Meerbeek, Stella Boess, Stefan Pauws, and Marieke Sonneveld. Robot vacuum cleaner personality and behavior. *IJSR*, 3(2):187–195, 2011.
- [7] Joseph L Jones. Robots at the tipping point: the road to irobot roomba. *IEEE Robotics & Automation Magazine*, 13(1):76–78, 2006.
- [8] Erwin Prassler, Arno Ritter, Christoph Schaeffer, and Paolo Fiorini. A short history of cleaning robots. *Autonomous Robots*, 9(3):211–226, 2000.
- [9] Bill Schackner. Cmu student wants to know how many are on campus; so far she’s up to 547. October 2010. [Online].
- [10] Ben Tribelhorn and Zachary Dodds. Evaluating the roomba: A low-cost, ubiquitous platform for robotics research and education. In *ICRA*, pages 1393–1399. IEEE, 2007.
- [11] Iwan Ulrich, Francesco Mondada, and J-D Nicoud. Autonomous vacuum cleaner. *Robotics and autonomous systems*, 19(3-4):233–245, 1997.



Cat track

Work-stealingsaving



34 **The Zero-color Theorem: An optimal poster design algorithm**

Michael Coblenz

35 **Cerebral genus: Dead duck or phoenix?**

Oscar I. Hernandez

The Zero-Color Theorem: An Optimal Poster Design Algorithm

Michael Coblenz
Computer Science Department
Carnegie Mellon University
Pittsburgh, PA
mcoblenz@cs.cmu.edu

ABSTRACT

Authors of academic posters frequently toil needlessly over their poster designs. In this paper, we show that the optimal number of colors for an academic poster is 0, significantly reducing the cost of poster production.

Keywords

Design; Algorithmic Presentation

1. INTRODUCTION

The problem of designing posters for academic conferences and other presentations has plagued numerous previous authors. The Cornell Center for Materials Research, for example, advises avoiding making one's poster resemble an abstract painting [2]. However, common poster designs fail to take into account the practicalities of how posters are actually used in the real world, leading researchers to develop severely suboptimal poster designs. In this paper, we show how to design the optimal poster, taking into account the actual usage pattern that is common of real academic posters. The paper culminates in a proof of the zero-color theorem, which gives optimal design guidance for academic posters. This is particularly useful for new members of the academic community, who may be tempted to use a suboptimal design for their posters.

2. ACADEMIC POSTER USE CASES

Traditionally, researchers have assumed that posters should be designed to impress potential viewers. This approach can be modeled as attempting to maximize B , the benefit of the poster presentation:

$$B = \sum_{i=1}^n B(v_i) \quad (1)$$

where v_i is the label of a particular viewer and B is a function that gives the benefit to the presenter of a particular

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

SIGBOVIK 2017 Pittsburgh, Pennsylvania, USA

© 2017 ACM. ISBN 123-4567-24-567/08/06...\$15.00

DOI: 10.475/123_4

viewer being impressed by the presenter's research. As $n \rightarrow \infty$ and $B(v) \rightarrow \infty$, it would seem that it would be optimal to expend an arbitrary amount of resources on poster creation.

Unfortunately, this approach fails to consider the overall poster lifecycle. The typical poster presentation lasts approximately an hour and a half [1]. However, after a poster is printed, it has an indefinite lifetime, potentially longer than that of the researcher who designed the poster. Some posters may last in excess of 1,000 years [4]. During the rest of the poster lifetime, the poster is typically stored inside a poster tube. As the poster lifetime l approaches ∞ , the fraction of the poster's lifetime spent in the display configuration approaches zero.

This observation leads to a novel result in poster design. Instead of optimizing for the nearly-nonexistent portion of the poster's lifetime in which it exists outside a poster tube, posters should be instead optimized for use *inside* poster tubes.

3. THE ZERO-COLOR THEOREM

In this section, we derive the optimal poster design of a poster for display inside a poster tube. Since the face of the poster is unobservable, one might initially assume that the content of the poster is inconsequential. Unfortunately, this is not the case. In fact, ink is quite expensive. As of 2011, ink cost approximately \$0.70 per mL, which is nearly twice the cost per mL of human blood [3]. Consider the number of colors, n used in the printing of a poster, where "used" means that a positive quantity of ink for the color has been expended in the poster's production. If $n > 0$, then the cost of ink, c , exceeds 0. But by assumption, no one can see the colors used in a poster while it is in a poster tube. Therefore, there is no benefit to using a nonzero number of colors. Since the cost of $n > 0$ is positive, we conclude that 0 is the optimal number of colors for a poster.

4. FUTURE WORK

Readers may observe that the theorem proved in this paper leads to a substantial simplification of the poster design process. Unfortunately, the problem of poster design is still complex and requires further study. Paper comes in an astounding array of different colors and finishes; though the result in this paper provides design guidance regarding number of *printed* colors, further work will be required to determine the optimal paper color and type.

5. CONCLUSIONS

Present at poster sessions with blank posters. Those who heed this advice will save their departments valuable funds and retain their access to the real point of poster sessions, namely snacks and drink tickets.

6. REFERENCES

- [1] This reasonable estimate sounds good to you too, right?
- [2] C. C. for Materials Research. Scientific poster design.
<http://hsp.berkeley.edu/sites/default/files/ScientificPosters.pdf>.
- [3] Visually. Ink costs more than human blood.
<http://visual.ly/ink-costs-more-human-blood>.
- [4] Wikipedia. Acid-free paper.
https://en.wikipedia.org/wiki/Acid-free_paper.



CONFIDENTIAL COMMITTEE MATERIALS

SIGBOVIK 2017 Paper Review

Paper 11: The Zero-Color Theorem An Optimal Poster Design Algorithm

Stefan Muller, Paper Collector

Rating: 0/0

Confidence: NaN

This paper introduces and proposes a proof of the zero-color theorem. I believe that the theorem, if true, would be an important advance in simplifying the poster design process. Unfortunately, I am not convinced by the result, which relies on the assumption that posters spend most of their life cycle in a poster tube. This, in turn, presupposes that graduate students can afford poster tubes, and do not simply attend poster sessions at their own institutions by rolling the poster up in a cylinder and shimmying a rubber band around it, the same state in which the poster will then find itself (either propped up in a corner or placed on a bookshelf). Keeping in mind this academic impoverishment factor, the theorem must consider two additional cases. If the poster is rolled with the printed side inward, the zero-color theorem may still hold. However, if the printed side is outward, some of the poster will be permanently displayed as it gathers dust. This means that the most important material on the poster must be squeezed into a space the width of the poster and the length $r\theta$, where r is the radius of the rolled-up poster and θ is the angle of the arc of poster which is visible. Within this strip, more than zero colors must be used. The paper should explore the optimal ink colors for this strip (most likely dark colors so that they will be visible once the poster has begun to fade).

Cerebral Genus

Dead Duck or Phoenix?

Oscar I. Hernandez

Division of Science, Mathematics, and Computing
Bard College at Simon's Rock
84 Alford Rd, Great Barrington, MA 01230
ohernandez13@simons-rock.edu

Abstract

I present an application of topology to the characterization of brain health.

1. Introduction

Recent studies suggest that one's life depends critically on the condition of their brain. We explore sufficient conditions to solve the following problem.

Problem 1. *Determine if your brain is unhealthy.*

The author decided to investigate this problem using topological methods when he learned[AW] that the word "topological" decomposes into "top" (head) + "o" (hole) + "logical" (smart).

2. Result

We use one of the most fundamental concepts of topology in order to solve the problem.

Definition 1. *The cerebral genus is the number of holes it has.*

Below is a necessary condition of a healthy brain.

Lemma 1. [OH13] *A healthy brain has a nonnegative genus.*

We use the well-known Positivity Lemma, stated above, in the main result.

Theorem 1. [OH13] *A positive cerebral genus g is a sign of an unhealthy brain.*

There is a safe proof, but I provide a constructive one below.

Proof. Carve a decently-sized hole in your brain, perhaps with a swift bullet. It is easy to see, say by asking your doctor, that you will die soon. \square

3. Conclusion

The reader is encouraged to remain open-minded and attempt to disprove the theorem. The author offers \$50 to the first reader who survives the attempt.

The health characterization of brains of genus 0 remains an open problem, as is the construction of jokes involving the phrases "pick your brain", "open minded", "closed minded".

References

- [AW] Aaron Williams. Personal communication. Mar 15, 2017.
- [OH13] Oscar Hernandez. Common Sense. 2013.

Track visualizations lifted from:

https://img.clipartfest.com/e433767404d82c4481bed8d20ed985e8_-amazoncom-bear-tracks-black-bear-tracks-clipart_522-218.jpeg

<http://i.ebayimg.com/images/i/321363254320-0-1/s-11000.jpg>

https://img.clipartfest.com/09f8577eafc2d3d111ca8915eb4d70c6_bird-tracks-clip-art-bird-tracks-clip-art_300-300.jpeg

http://www.ultimatefieldguide.com/Vervet_Monkey_Track.gif

http://static.dusupply.com/media/catalog/product/cache/1/image/9df78eab33525d08d6e5fb8d27136e95/d/o/dog_tracks2.jpg

<http://extension.missouri.edu/explore/images/g09452track01.jpg>

http://www.biokids.umich.edu/images/signs/tracks/chipmunktracks_drawing_thum.jpg

<http://www.polyvore.com/cgi/img-thing?.out=jpg&size=1&tid=11217414>

<http://www.ankn.uaf.edu/curriculum/units/images/2moose2.GIF>

<http://nyfalls.com/dev/wp-content/uploads/2013/04/raccoon-tracks1.gif>

<http://www.clipartbest.com/cliparts/ycog9G/ycog9Gzgi.gif>