



<http://researchspace.auckland.ac.nz>

ResearchSpace@Auckland

Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

To request permissions please use the Feedback form on our webpage.

<http://researchspace.auckland.ac.nz/feedback>

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form.

Approximability and Computational Feasibility of Dodgson's Rule

Supervisors: Dr A. Slinko, Dr G. Pritchard

John C. McCabe-Dansted

June 7, 2006



<http://researchspace.auckland.ac.nz>

ResearchSpace@Auckland

Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

To request permissions please use the Feedback form on our webpage.

<http://researchspace.auckland.ac.nz/feedback>

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form.

Abstract

Condorcet¹ proposed that a winner of an election is not legitimate unless a majority of the population prefer that alternative to all other alternatives. However such a winner does not always exist. A number of voting rules have been proposed which select the Condorcet winner if it exists, and otherwise selects an alternative that is in some sense closest to being a Condorcet Winner; a prime example is the rule proposed by Dodgson²(1876).

Unfortunately, Bartholdi et al. (1989) proved that finding the Dodgson winner is an NP-hard problem. Hemaspaandra et al. (1997) refined this result by proving that it is Θ_2^p -complete and hence is not NP-complete unless $\Theta_2^p = \text{NP}$. For this reason, we investigate the asymptotic behaviour of approximations to the Dodgson rule as the number of agents gets large.

Under the assumption that all votes are independent and equiprobable, the probability that the Tideman (1987) approximation picks the Dodgson winner does asymptotically converge to 1, but not exponentially fast. We propose a new approximation that does exhibit exponential convergence, and we can quickly verify that it has chosen the Dodgson winner; this allows us to choose the true Dodgson winner with $\mathcal{O}(\ln n)$ expected running time for a fixed number of alternatives m and n agents.

McGarvey (1953) proved that all tournaments are the majority relations for some society. We have proved a generalisation of this theorem for weighted tournaments. We find that this generalisation is useful for simplifying proofs relating to rules which use the weighted majority relation.

Bartholdi et al. (1989) found that we can calculate the Dodgson Score using an ILP that requires no more than $m!m$ variables, we present an improved ILP that requires less than $(m - 1)!e$ variables ($e \approx 2.71$). We discover that we can solve this ILP in $\mathcal{O}(\ln n)$ arithmetic operations of $\mathcal{O}(\ln n)$ bits in size. Relaxing the integer constraints results in a new polynomial time rule. In 43 million simulations this new rule failed to pick the

¹Marie-Jean-Antoine-Nicolas de Caritat, Marquis de Condorcet. (1743–1794)

²Charles Lutwidge Dodgson (1832–1898), better known as Lewis Carroll.

Dodgson winner only once, and only given many (25) alternatives. Unlike the Dodgson rule, this rule can break ties in favor of alternatives that are in some sense fractionally better.

We show that Dodgson Score admits no constant error approximation unless $P=NP$, and admits no Polynomial Time Approximation Scheme (PTAS) for Dodgson Score unless $W[2]=FPT$.



Acknowledgements

I would first of all like to thank my sister, Kim Dansted, for her suggestions for improving the readability of my introduction and for her general support and assistance while I was going through the grueling process of turning a draft into a legible Thesis.

Dr Arkadii Slinko, my supervisor for much assistance with proof reading. Also, unlike most other Lecturers, he helped guide me through my developmental years, due to his work with the pre-tertiary International Mathematical Olympiad training camps. Indeed, this was why I sought him out as a supervisor.

Dr Geoff Pritchard, my co-supervisor, for assistance with the statistics. Thanks to Pritchard, I have a more rigorous proof that Tideman's approximation converges to Dodgson's rule, which makes proper use of the Multivariate Central Limit Theorem.



Contents

List of Tables	ix
1 Introduction	1
1.1 Introduction	1
1.1.1 Borda’s Objection to the Condorcet proposal	2
1.1.2 Impossibility Theorems	4
1.1.3 Complexity Classes for Algorithms	5
1.1.4 Impracticality Theorems	10
1.1.5 Simplifying Assumptions in Voting Theory	11
1.1.6 Tideman’s Approximation to the Dodgson Rule	11
1.1.7 Our New Approximation, Dodgson Quick.	13
1.1.8 Linear Programs and Integer Linear Programs	16
1.1.9 Our New Approximation, Dodgson Relaxed & Rounded.	18
1.1.10 A Generalisation of the McGarvey Theorem	19
1.2 Social Choice Functions.	19
1.2.1 Fishburn’s Classification System for Voting Rules	20
1.2.2 Advantages	22
1.2.3 Condorcet Winner	22
1.2.4 Scores	22
1.2.5 Impartial Culture Assumption	24
1.2.6 Pólya-Eggenberger Urn Model	25
1.3 Summary	25
2 A McGarvey Theorem for Weighted Tournaments	29
3 Simple Rules that Approximate the Dodgson Rule	35
3.1 Dodgson Quick, A New Approximation	35

Contents

3.2	Tideman's Rule	41
3.3	Numerical Results	53
3.3.1	Asymptotic Behaviour of Simpson's Rule	56
4	Formulation of Dodgson's Rule as an Integer Linear Program	59
4.1	Discussion of Variables	60
4.2	Preliminary Definitions	63
4.3	Definition of Integer Linear Program	64
4.4	Number of Variables	67
4.5	Size of the Linear Program	68
5	Dodgson Relaxed & Rounded	71
5.1	Definition of Dodgson Relaxed Score	72
5.2	Complexity	77
6	Feasibility of Dodgson's Rule	79
6.1	Theoretical Worst-Case Results	80
6.2	Approximability Classes	81
6.3	Empirical Performance of Dodgson's Rule.	88
7	Conclusion	91
7.1	Methodological Issues	93
7.2	Further Work	94
A	Preliminary Mathematics	97
A.1	Probability Space	97
A.2	Binomial Distribution	99
A.3	Multinomial Distribution	101
A.4	Multivariate Normal Distribution	106
A.5	Multisets	107
B	Code and Output	109
B.1	Asymptotic Simpson's Rule	109
B.1.1	asypm.sh — wrapper script	109
B.1.2	asypm.c	110
B.1.3	Output	115
B.2	Dodgson Quick vs Tideman vs Simpson	116

B.2.1	SiTiDQ.sh — wrapper script	116
B.2.2	SiTiDQ.c	116
B.2.3	Output	122
B.3	Exact Dodgson Algorithm	125
	Reference List	143
	Index	147

Contents



List of Tables

3.1	Number of Occurrences per 1000 Elections with 5 Alternatives that the Dodgson Winner was Not Chosen	53
3.2	Number of Occurrences per 1000 Elections with 5 Alternatives that the Set of Tied Dodgson Winners was Not Chosen	54
3.3	Frequency that the DQ-Winner is the Dodgson Winner	55
3.4	Frequency that the Tideman Winner is the Dodgson winner	55
3.5	Frequency that the Simpson Winner is the Dodgson Winner	55
3.6	The Limit of the Number of Occurrences per 1000 Elections that the Simpson Winner is Not the Dodgson Winner, as $n \rightarrow \infty$	56
5.1	Example of Dodgson Score of d Differing from the Relaxed Score	73
5.2	Occurrences out of 80,000 that the Dodgson Relaxed Winner Differed from the Dodgson Winner after Tie-Breaking.	75
5.3	Occurrences out of 80,000 that the Set of Tied Dodgson Relaxed Winners Differed from the Set of Tied Dodgson Winners.	75
5.4	Occurrences out of 80,000 that the Set of Tied Dodgson Relaxed Winners were not a Subset of the Set of Tied Dodgson Winners.	76
6.1	Time in Milliseconds to Compute the Dodgson Winner (#Alter/#Voter,b=0) 88	
6.2	Time in Milliseconds to Compute the Dodgson Winner (5 alternatives,b=0)	88
6.3	CPU Time in Seconds to Calculate the Dodgson Winner in Impartial Culture for Square Profile ($n = m = s$).	88

LIST OF TABLES

Notations

- x^+ : 0 if $x < 0$, x otherwise.
- $\text{sgn}(x)$: The sign of x , 1 if $x > 0$, -1 if $x < 0$, 0 if $x = 0$.
- $\lfloor x \rfloor$: Floor of x , the largest integer that is less than or equal to x .
- $\lceil x \rceil$: Ceiling of x , the smallest integer that is greater than or equal to x .
- $|x|$: Absolute value of x , that is x if $x > 0$, $-x$ otherwise.
- \sum_i : Summation over each i .
- \prod_i : Product over each i .
- $m!$: The factorial of m , i.e. $m! = (1)(2)(3) \cdots (m) = \prod_{i=1}^m i$.
- $A \wedge B$: A and B , e.g. A and B will occur.
- $A \vee B$: A or B , e.g. A or B will occur.
- $\neg A$: The negation of A , i.e. the statement “ A is false”.
- $P(E)$: The probability of event E , a real number in $[0, 1]$.
- $P(A|B)$: The probability of event A will occur given B has occurred or will occur.
- $\text{var}(X)$: variance of random variable X .
- $E[X]$: mean of random variable X .
- $A \xrightarrow{D} B$: A converges in distribution to B .
- $N(\mu, \Omega)$: A multivariate normal distribution with means of μ and matrix of correlations Ω .

- $\mathcal{L}(A)$: The set of all linear orders on the set of alternatives A .
- \mathbb{Z} : The set of integers.
- $2\mathbb{Z}$: The set of even numbers.
- \mathbb{N} : The set $\{1, 2, 3, \dots\}$ of natural numbers. (also known as positive integers)
- \mathbb{R} : The set of real numbers, e.g. $0, 1, \pi, \sqrt{2}$.
- \mathbb{R}^+ : The set of positive real numbers.
- \mathbb{R}^k : The set of real valued k -dimensional vectors.
- $[x, y]$: The range of real numbers from x to y inclusive.
- (x, y) : The range of real numbers from x to y , including x but not y .
- avb : Alternative a is ranked before b in linear order \mathbf{v} .
- $\bar{\mathbf{v}}$: The opposite of a linear order \mathbf{v} , i.e. $avb \Leftrightarrow b\bar{\mathbf{v}}a$.
- n_{ba} : The number of linear orders in our profile \mathcal{P} where alternative b is ranked above alternative a .
- $\text{adv}(b, a)$: The advantage of b over a , defined as $(n_{ba} - n_{ab})^+$.
- $\mathbf{1}_k$: A k -dimensional vector with all subscripts equaling 1.
- M^T : The transpose of M , i.e. $M_{ij} \equiv (M^T)_{ji}$.
- $\#(S)$: The cardinality of the set S , i.e. the number of elements in S .
- $F(x) \in \mathcal{O}(G(x))$: Function F is of order G , i.e. there exists $N \in \mathbb{N}$ and $c > 0$ such that for all n greater than N , $F(x) \leq cG(x)$.
- $\mathbf{x} \leq \mathbf{y}$: vector $\mathbf{x} = (x_1, x_2, \dots, x_n)$ is less than vector $\mathbf{y} = (y_1, y_2, \dots, y_n)$. That is, $x_i \leq y_i$ for all i in $\{1, 2, \dots, n\}$.
- $\mathbf{x} \geq 0$: vector $\mathbf{x} = (x_1, x_2, \dots, x_n)$ is positive. That is, $x_i \geq 0$ for all i in $\{1, 2, \dots, n\}$.
- $A \subseteq B$: A is a subset/submultiset of B .
- x^n : n instances of x , e.g. $\{a, b^2, c\}$ is a multi set with one instance of a and c and two instances of b .
- $\ln x$: The natural log of x , i.e. $\log_e x$.