Dodgson Winner Problem

Dhamma Kimpara

August 5, 2020

1 Introduction

Suppose you are a theoretical computer scientist, wandering in the wilds of theory-land and you come upon a complexity class. It has not been seen before. It seems intuitive and natural. So you decide to submit it to the complexity zoo [Aaronson et al., 2005]. But the worry is whether this class will turn out to capture the complexity of important real-world problems. In other words, is anyone or any problems going to come around to your new section of the zoo and join your party? After all, there are 545 classes and counting in the zoo!

This was the case for the class Θ_2^pin the mid-1990s. Θ_2^por $\mathsf{P}_{||}^{\mathsf{NP}}$, the class of languages decidable by a polynomial time Turing machine with parallel access to NP, was introduced in 1983 by [Papadimitriou and Zachos, 1982]. The location of this class in the polynomial hierarchy is:

By the mid 1990s, the theoretical importance of $\Theta_2^p$was recognized in complexity theory. Klaus W. Wagner established half a dozen characterizations of Θ_2^p[Wagner, 1990], several complete problems and a toolkit for establishing Θ_2^p-hardness. Hemachandra [1987] and Köbler, Johannes et al. [1987] showed that $\Theta_2^p$was equivalent to the class of problems that can be solved by $\mathcal{O}(\log(n))$ sequential Turing queries to NP. Furthermore, if NPcontains some Θ_2^p-hard problem, then the polynomial hierarchy collapses to NP. However, these connections and complete problems lived in the pure theory section of the zoo and did not yet have the appeal of problems from "real world" settings.

Along came the Dodgson winner problem, invented in 1876 by Charles Lutwidge Dodgson, better known under the pen name of Lewis Carroll. Hemaspaandra et al. [1999] proved that this problem was Θ_2^p-complete. This was the first "real-world" problem proven complete for the class Θ_2^p.

Dodgson's election system takes the following form. An election is a finite number of voters, who each cast a linear order over a common finite set of candidates. Note that linear orders are "tie-free". The winner is determined by whichever candidate is closest to being a Cordorcet winner, a criteria used by other election systems. A Condorcet winner is a candidate a who for every other candidate b, is preferred to b by strictly more than half of the voters. We naturally want election systems to be Condorcet-consistent, i.e. the system has the property where if a is a Condorcet winner, a is the one and only winner in the election. Dodgson's election system is Condorcet-consistent [Brandt et al., 2016]

The winner(s) in a Dodgson election is defined as the candidate(s) who are the "closest" to being Condorcet winners. The winners are the candidates that have the lowest Dodgson score. The Dodgson score of a candidate a is the smallest number of sequential exchanges of adjacent candidates in preference orders such that after those exchanges a is a Cordorcet winner.

Note that it is remarkable that we find Θ_2^p-complete problems that were defined 100 years before complexity theory itself existed. Dodgson winner is also extremely natural when compared with other complete problems in this class such as determining if the maximum size clique in a graph is of odd size (odd-max-clique). In this project we will use Comp-SAT, which was recently shown to be complete for this class [Lukasiewicz and Malizia, 2017]

The rest of this project presents the theory of and a practical algorithm for the Dodgson winner problem. We present slightly modified proof of completeness of the Dodgson winner problem based on new results for the class Θ_2^p. Then we present, implement, and examine a heuristic algorithm that is self-knowingly correct for most practical instances of the problem.

2 Preliminaries

2.1 Problem Setting

Definition 1. Dodgson Triple
A triple, $\langle C, c, V\rangle$, where C is the set of candidates $1, \ldots, m$, a candidate $c \in C$, and a set of n strict (ie. irreflexive and anti-symmetric) preference orders, one per voter, over all candidates in C.

Definition 2. Condorcet Winner
In an election, with a set of candidates C and n votes or strict preference orders V, a candidate $a \in C$ is a Condorcet Winner if for every other candidate b, $a \succ b$ by strictly more than half of the voters.

Definition 3. Dodgson Score
First define a switch as an exchange of two adjacent preferences in the preference order of one voter. Then, the Dodgson Score of a candidate is the smallest number of sequential switches needed to make the candidate a Condorcet winner. The Dodgson Score of any Cordorcet Winner is 0. We denote the Dodgson score of a Dodgson triple as $Score(\langle C, c, V\rangle)$

Decision Problem. DODGSONSCORE
Instance: $k \in \mathbb{N}$. A Dodgson Election and Candidate $\langle C, c, V\rangle$.
Decide: Is the Dodgson Score of candidate c less than or equal to k?

Decision Problem. DODGSONWINNER
Instance: A Dodgson Election and Candidate $\langle C, c, V\rangle$
Decide: Is c a winner of the election? In other words, does c have the minimum Dodgson Score in the election?

Decision Problem. COMP-SAT [Lukasiewicz and Malizia, 2017]
Instance: A pair $\langle A, B\rangle$ of sets of 3CNF formulas.
Decide: Is the number of satisfiable formulas in A greater than the number of satisfiable formulas in B.

2.2 Complexity Classes and Definitions

Definition 4. Θ_2^p:
The class of problems solvable with polynomial-time parallel access to an NPoracle. This is equivalent to $\mathcal{O}(\log(n))$ sequential queries to an NPoracle [Hemachandra, 1987, Köbler, Johannes et al., 1987].

Theorem 1. [Lukasiewicz and Malizia, 2017] COMP-SAT is Θ_2^p-complete.

3 The complexity of the Dodgson Winner Problem

Our results largely follow the proof in Hemaspaandra et al. [1999], except instead of using a technical lemma by Wagner [1990], we instead begin the reduction with the recent result that COMP-SAT is Θ_2^p-complete [Lukasiewicz and Malizia, 2017].

We first state the main result of this section and a big-picture corollary concerning the polynomial hierarchy.

Theorem 2. DODGSONWINNER is Θ_2^p-complete

It follows that though DODGSONWINNER is NP-hard [Bartholdi et al., 1989], it cannot be NP-complete unless the polynomial hierarchy collapses.

Corollary 1. If DODGSONWINNER is NP-complete then $\mathsf{PH} = \mathsf{NP}$.

3.1 Outline of Proof of Theorem 2

First we briefly outline the reduction from Comp-SAT to DodgsonWinner. Let us start with a COMP-SAT instance $\langle A, B\rangle$, for each 3CNF F formula in A and B, reduce F into the corresponding instance of THREEDIMENSIONALMATCHING (3DM), which is possible because 3DM is NP-complete. Now, for each 3DM problem, we reduce it into an instance of DodgsonScore. Using a merger lemma, we will merge these into 2 elections, one for A and one for B, thus an instance of 2ELECTIONRANKING (2ER). Then, 2ER will be reduced to DodgsonWinner.

A reference diagram is provided:

$$\text{COMP-SAT} \leq_m^p \text{COMP-3DM} \leq_m^p \text{COMP-DODGSONSCORE} \leq_m^p \text{2ER} \leq_m^p \text{DODGSONWINNER}$$

Now we need to show membership in Θ_2^p.

Theorem 3. DODGSONWINNER $\in \Theta_2^p$.

Proof: Ask in parallel all the DODGSONSCORE queries, one for each candidate in C. Each query is an NPquery since DODGSONSCORE is NP-complete by the second reduction outlined above (lemma 1). We now have the exact Dodgson Score for each candidate. It is easy to decide whether or not the given candidate c ties-or-defeats all other candidates in the election (find the max of the scores and compare the c's score to the max $\sim \mathcal{O}(n)$).

3.2 Outline of relevant lemmas

Lemma 1. There exists an NP-complete problem A and a polynomial-time computable function f that reduces A to DODGSONSCORE in such a way that, $\forall x \in \Sigma^*, f(x) = \langle\langle C, c, V\rangle, k\rangle$ is an instance of DODGSONSCORE with an odd number of voters and

1. if $x \in A$ then $Score(\langle C, c, V\rangle) = k$, and
2. if $x \notin A$ then $Score(\langle C, c, V\rangle) = k + 1$.

Lemma 2. There exists a polynomial-time computable function $DodgsonSum$ such that $\forall k$ and for all $\langle C_1, c_1, V_1\rangle, \langle C_2, c_2, V_2\rangle, \ldots, \langle C_k, c_k, V_k\rangle$ satisfying $||V_j||$ is odd for all j, it holds that

$$DodgsonSum(\langle\langle C_1, c_1, V_1\rangle, \langle C_2, c_2, V_2\rangle, \ldots, \langle C_k, c_k, V_k\rangle\rangle)$$

is a Dodgson triple having an odd number of voters and such that

$$\sum_j Score(\langle C_j, c_j, V_j\rangle) = Score(DodgsonSum(\langle\langle C_1, c_1, V_1\rangle, \langle C_2, c_2, V_2\rangle, \ldots, \langle C_k, c_k, V_k\rangle\rangle))$$

Theorem 1, Lemma 1 and Lemma 2, establish the Θ_2^p-hardness of a problem related to DODGSONWINNER. We now define this decision problem:

Decision Problem. TWOELECTIONRANKING (2ER)

Instance: A pair of Dodgson triples $\langle\langle C, c, V\rangle, \langle D, d, W\rangle\rangle$ both having an odd number of voters such that $c \neq d$.

Decide: Is $Score(\langle C, c, V\rangle) \leq Score(\langle D, d, W\rangle)$?

Lemma 3. 2ER is Θ_2^p-hard.

We now need to make the results so far applicable to DODGSONWINNER, so we need another merger lemma to merge two elections into a single election.

Lemma 4. There exists a polynomial-time computable function $Merge$ such that, for all Dodgson triples, $\langle C, c, V\rangle$ and $\langle D, d, W\rangle$ for which $c \neq d$ and both having an odd number of voters, there exist $\hat{C}$ and $\hat{V}$ such that

1. $Merge(\langle C, c, V\rangle, \langle D, d, W\rangle)$ is an instance of DodgsonWinner,
2. $Merge(\langle C, c, V\rangle, \langle D, d, W\rangle) = \langle \hat{C}, c, \hat{V}\rangle$,
3. $Score(\langle \hat{C}, c, \hat{V}\rangle) = Score(\langle C, c, V\rangle) + 1$,
4. $Score(\langle \hat{C}, d, \hat{V}\rangle) = Score(\langle D, d, W\rangle) + 1$ and,
5. for each $e \in \hat{C} \setminus \{c, d\}$, $Score(\langle \hat{C}, c, \hat{V}\rangle) < Score(\langle \hat{C}, e, \hat{V}\rangle)$

3.3 Proof of select lemmas

Proof: Lemma 1 or 3DM $\leq_m^p$ DodgsonScore
Our reduction differs from Bartholdi et al. [1989] in that our reduction has additional properties that are required by the lemma. We will reduce from ThreeDimensionalMatching to DodgsonScore:

Decision Problem. ThreeDimensionalMatching (3DM)

Input: Sets M, W, X, Y, where $M \subseteq W \times X \times Y$ and W, X, Y are disjoint, nonempty sets having the same number of elements.

Decide: Does M contain a matching, i.e. a subset $M' \subseteq M$ such that $M' = W$ and no two elements of M' agree in any coordinate?

Proof: Lemma 3
We will reduce from Comp-SAT to 2ER. Let $\langle A, B\rangle$ be a Comp-SAT instance.

For each 3CNF formula $x \in A$ or B, reduce x into the corresponding 3DM instance x' and add x' to A' or B' if $x \in A$ or $x \in B$, respectively. In effect we are reducing Comp-SAT to an instance of Comp-3DM, $\langle A', B'\rangle$. It is easy to see that solving $\langle A', B'\rangle$ solves $\langle A, B\rangle$.

Now we perform a similar reduction from Comp-3DM to Comp-DodgsonScore. For each $x' \in A'$ or B', use the function in Lemma 1 to reduce x' into the corresponding DodgsonScore instance $\langle C, c, V\rangle$ and add $\langle C, c, V\rangle$ to A'' or B'' if $x' \in A'$ or $x' \in B'$, respectively. It is similarly easy to see that solving $\langle A'', B''\rangle$ solves $\langle A', B'\rangle$. If $x' \in$ Comp-3DM then by Lemma 1 $Score(\langle C, c, V\rangle) = k$ and thus DodgsonScore on $\langle C, c, V\rangle$ returns yes. If $x' \notin$ Comp-3DM then $Score(\langle C, c, V\rangle) = k + 1$ and DodgsonScore on $\langle C, c, V\rangle$ returns no.

Now we reduce from Comp-DodgsonScore to 2ER using Lemma 2. First note that the direction of the inequality of the decision problem changes in this reduction by the nature of Lemma 1. Now to begin the reduction, we merge all the Dodgson elections in A'', B'' into $\langle C, c, V\rangle$ and $\langle D, d, W\rangle$, respectively. This is done using the $DodgsonSum$ function in 2. For example if $A'' = \langle C_1, c_1, V_1\rangle, \langle C_2, c_2, V_2\rangle, \ldots, \langle C_k, c_k, V_k\rangle$ then

$$\langle C, c, V\rangle = DodgsonSum(\langle\langle C_1, c_1, V_1\rangle, \langle C_2, c_2, V_2\rangle, \ldots, \langle C_k, c_k, V_k\rangle\rangle)$$

and

$$\sum_j Score(\langle C_j, c_j, V_j\rangle) = Score(\langle C, c, V\rangle).$$

Now $\langle A'', B''\rangle$ being true is equivalent to the number of satisfied DodgsonScore instances in A'' being greater than that of B''. This again is equivalent to the sum of the Dodgson scores of A'' being less than that of B''. Thus using the reduction of $\langle A'', B''\rangle$ outlined above, If $Score(\langle C, c, V\rangle) \leq Score(\langle D, d, W\rangle)$ is true, then $\langle A'', B''\rangle$ is true and similarly for the false case. Hence we have shown a reduction from Comp-DodgsonScore to 2ER.

Combining the many-one reductions above, we have shown that:

$$\textsc{Comp-SAT} \leq_m^p \textsc{Comp-3DM} \leq_m^p \textsc{Comp-DodgsonScore} \leq_m^p \text{2ER}.$$

So by Theorem 1, 2ER is Θ_2^p-hard.

3.4 Proof that DodgsonWinner is Θ_2^p-hard

Proof: Theorem 2
We prove by showing that 2ER $\leq_m^p$ DodgsonWinner.

4 Practical Greedy Algorithm

5 Fixed Parameter Tractability

Proposition 1. 3.10 d winner is FPT

6 Experiments

7 Conclusion

References

Scott Aaronson, Greg Kuperberg, and Christopher Granade. The complexity zoo, 2005.

J. Bartholdi, C. A. Tovey, and M. A. Trick. Voting schemes for which it can be difficult to tell who won the election. *Social Choice and Welfare*, 6(2):157–165, 1989. ISSN 01761714, 1432217X. URL http://www.jstor.org/stable/41105913.

Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D Procaccia. *Handbook of computational social choice*. Cambridge University Press, 2016.

L. A. Hemachandra. The strong exponential hierarchy collapses. In *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing*, STOC '87, page 110–122, New York, NY, USA, 1987. Association for Computing Machinery. ISBN 0897912217. doi: 10.1145/28395.28408. URL https://doi.org/10.1145/28395.28408.

Edith Hemaspaandra, Lane A. Hemaspaandra, and Jörg Rothe. Exact analysis of dodgson elections: Lewis carroll's 1876 voting system is complete for parallel access to NP. *CoRR*, cs.CC/9907036, 1999. URL https://arxiv.org/abs/cs/9907036.

Köbler, Johannes, Schöning, Uwe, and Wagner, Klaus W. The difference and truth-table hierarchies for np. *RAIRO-Theor. Inf. Appl.*, 21(4):419–435, 1987. doi: 10.1051/ita/1987210404191. URL https://doi.org/10.1051/ita/1987210404191.

Thomas Lukasiewicz and Enrico Malizia. A novel characterization of the complexity class thetakp based on counting and comparison. *Theoretical Computer Science*, 694: 21 – 33, 2017. ISSN 0304-3975. doi: https://doi.org/10.1016/j.tcs.2017.06.023. URL http://www.sciencedirect.com/science/article/pii/S0304397517305352.

Christos H. Papadimitriou and Stathis K. Zachos. Two remarks on the power of counting. In Armin B. Cremers and Hans-Peter Kriegel, editors, *Theoretical Computer Science*, pages 269–275, Berlin, Heidelberg, 1982. Springer Berlin Heidelberg. ISBN 978-3-540-39421-1.

Klaus W Wagner. Bounded query classes. *SIAM Journal on Computing*, 19(5):833–846, 1990.