JOURNAL OF INFORMATION SECURITY AND APPLICATIONS 19 (2014) 115–129

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jisa

Verifying voting schemes

Bernhard Beckert [a,*], Rajeev Goré [b], Carsten Schürmann [c,**], Thorsten Bormer [a], Jian Wang [c]

[a] Karlsruhe Institute of Technology, Am Fasanengarten 5, 76131 Karlsruhe, Germany
[b] Research School of Computer Science, The Australian National University, Australia
[c] IT University of Copenhagen, Rued Langgaards Vej 7, 2300 Copenhagen S, Denmark

ARTICLE INFO

Article history:
Available online 4 July 2014

Keywords:
Voting schemes
Verification
Bounded model checking
Single transferable vote
SMT solvers

ABSTRACT

The possibility to use computers for counting ballots allows us to design new voting schemes that are arguably fairer than existing schemes designed for hand-counting. We argue that formal methods can and should be used to ensure that such schemes behave as intended and conform to the desired democratic properties. Specifically, we define two semantic criteria for single transferable vote (STV) schemes, formulated in first-order logic over the theories of arrays and integers, and show how bounded model-checking and SMT solvers can be used to check whether these criteria are met. As a case study, we then analyse an existing voting scheme for electing the board of trustees for a major international conference and discuss its deficiencies.

1. Introduction

The goal of any social choice function is to compute an "optimal" choice from a given set of preferences. Voting schemes in elections are a prime example of such choice functions as they compute a seat distribution from a set of preferences recorded on ballots. By *voting scheme* we refer to the method for counting ballots and computing who won – as opposed to an actual computer implementation of such a scheme or a scheme describing the process of casting votes via computer. The difficulty in designing preferential voting schemes is that the optimisation criteria are not only multi-dimensional, but multi-dimensional on more than one level. On one level, we want to satisfy each voter, so each voter is a dimension. On a higher level, there are desirable global criteria such as "majority rule" and "minority protection" that are at least partly inconsistent with each other. It is well-known that "optimising" such theoretical voting schemes along one dimension may cause them to become "sub-optimal" along another.

This observation is not new and voting specialists have proposed a series of mathematical criteria (Brandt et al., 2012) that can be used to compare various voting schemes with one another. A classic example is the notion of a Condorcet winner, defined as the candidate who wins against *each* other candidate in a one-on-one contest. Such a winner exists provided that there is no cycle in the one-to-one contest relation. A voting scheme is said to satisfy the Condorcet criterion if the Condorcet winner is guaranteed to be elected when such a winner exists. Another is the *monotonicity criterion* which

* *Corresponding author*. Tel.: +49 72160844025.

** *Corresponding author*.

E-mail addresses: beckert@kit.edu (B. Beckert), Rajeev.Gore@anu.edu.au (R. Goré), carsten@demtech.dk, carsten@itu.dk (C. Schürmann), bormer@kit.edu (T. Bormer), jwan@itu.dk (J. Wang).

http://dx.doi.org/10.1016/j.jisa.2014.04.005
2214-2126/© 2014 Elsevier Ltd. All rights reserved.

requires that a candidate who wins a contest will also win if the ballots were changed uniformly to rank that candidate higher.

In practice, theoretical voting schemes are often simplified in many ways when used in real-world elections, typically to reduce their complexity to allow counting by hand. Such practical schemes may not satisfy general properties such as the Condorcet criterion simply because it is intractable to compute the Condorcet winner by hand, but they may satisfy some weaker version of "optimality" that is specific to that particular scheme. It may even happen that one among the optimal winners is chosen at random (Brams and Sanver, 2003) (as allowed by the Australian Capital Territory's Hare-Clark Method) or that someone other than the optimal winner is elected.

Voting schemes also evolve over time – for national elections in the large, and local elections, union elections, share holder elections, and board of trustee elections in the small. Incremental changes to the electoral system, the tallying process and the related algorithms challenge the common understanding about what the voting scheme actually does. For example, since 1969 some local elections in New Zealand adopted Meeks' method (Hill et al., 1987), which is a voting scheme for preferential voting that uses fractional weightings in its computations and is too complex to count by hand. This also required an adjustment of understanding about who will now be elected. In general, it is often not clear whether changes to the electoral system improve or worsen the overall quality of a voting scheme with regard to the various dimensions of optimisation. Changes to the electoral system in Germany, for example, have created paradoxical situations where more votes for a party translate into fewer seats and fewer votes into more seats, and have prompted Germany's Supreme Court to intervene at several occasions (see, e.g., Bundesverfassungsgericht, 2008).

Many jurisdictions around the world are now using computers to count ballots according to traditional voting schemes. Using computers to count ballots opens up the possibility to use voting schemes which really are optimised along multiple dimensions, while retaining global *desiderata* such as the Condorcet criterion. The inherent complexity of counting ballots according to such schemes means that it may no longer be possible to "verify" the result by hand-counting, even when the number of ballots is small. It is therefore important to imbue these schemes with the trust accorded to existing schemes. Note that our focus is on trust in the voting scheme, not trust in the computer-based process for casting votes.

One way to engender trust in such complex yet "fairer" voting schemes is to specify the *desiderata* when the scheme is being designed, and then formally check that the scheme meets these criteria before proposing changes to the legislation to enact the scheme. Such formal analyses could contribute significant unbiased information into the political discussions that typically involve such legislative changes and also assure voters that the changes will not create paradoxical situations as described above.

Formal analysis, however, is only practicable when we possess formal specifications of the voting scheme. We argue that it is important to give declarative specifications of the properties of a voting scheme for two reasons: (1) For understanding their properties and how they change during the evolution process, so that improving a scheme in one aspect does not by accident introduce flaws w.r.t. other aspects. (2) For checking the correctness of the scheme from both an algorithmic and implementation perspective. We also argue that general criteria are not sufficient and criteria are needed that are tailor-made for specific (classes of) voting schemes.

The properties in question are difficult to state, to formalise, to understand, to analyse, and to describe declaratively (as opposed to algorithmically) because: the final voting scheme may have to compromise between the conflicting demands of multiple individual desirable properties; the voting scheme may evolve and we may have to revisit these desiderata; even when the properties can be made mathematically precise, the resulting mathematical statement cannot serve as a specification if the electoral law defines a voting scheme that does not (always) compute the optimal solution.

Contributions. Here, we show that seemingly innocuous revisions to a voting scheme can have serious implications on the desired properties of the system and how analysis techniques employing Satisfiability Modulo Theories (SMT) solvers (De Moura and Bjørner, 2008) can be used to discover them. As a running example, we use the preferential voting scheme single transferable vote (STV) that is used in large national elections world-wide, but also for smaller professional elections.

In Sec. 3, we define two tailor-made criteria to establish the desired properties of the voting scheme. Both criteria are formulated using first-order logic and are amenable for bounded model checking with Z3, which is the tool of choice for our formal analysis (Sec. 4). Besides the experiments, we also discuss advantages and disadvantages of different verification techniques. Subsequently, we discuss (Sec. 5) a particularly interesting variant of the Single Transferrable Vote Algorithm (CADE-STV) for the board of trustees of the International Conference on Automated Deduction (CADE). We explain its oddities and differences to standard STV, and give a historical account of the conception and the stepwise refinement of the algorithm. This paper extends our previous work on the specification and verification of voting schemes (Beckert et al., 2013a) and also our system description of a bounded model checking system for analysing voting schemes and its application to CADE-STV (Beckert et al., 2013b).

Related work. Voting schemes have been investigated by social choice theorists for many decades. These tend to be mathematical analyses which prove various (relative) properties of different voting schemes: see (Pacuit and Zalta, 2012; Arrow, 1950). Such work tends to concentrate on what we have referred to as theoretical schemes and is often couched in terms of a formal theorem and its proof in natural language.

There is also a significant body of research on various properties of vote-casting schemes, particular security properties (Sun et al., 2012).

There does not appear to be much existing work on the formal analysis of voting schemes using methods and tools