Election Validity and Openness

Todd Davies
Symbolic Systems Program
Stanford University
(davies@stanford.edu)

10th International Meeting of the
Society for Social Choice and Welfare
Moscow, Russia,
July 22, 2010

(comments welcome
– incomplete and sketchy –
please do not quote without permission)

slides at http://bit.ly/c6p8gF

I will consider the conditions under which an election outcome can be considered an accurate (or valid) reflection of the preferences and intentions of legitimate voters, applying epistemic logic and common knowledge to the tradeoff between election validity and the openness of the voter and vote lists.

“The task of election administrators is to convince the losers that they lost.”
– David Dill

I. Model

We begin with a model that assumes the following (assume function values are nil outside of their domains):

- An overall population set P.
- A set of eligible voters E.
- A set of reported voters V of size M.
- A set of alternatives $A=\{0,1\}$.
- The actual voters $v_a: V \rightarrow P$.
- Intent $i: V \rightarrow A$.
- Selection $s: V \rightarrow A$.
- Marked Ballot $b_m: V \rightarrow A$.
- Counted Ballot $b_c: V \rightarrow A$.
- Tally $t: V \rightarrow A$.
- Reported Result $R \in \{0,\ldots,M.\}$

II. Epistemic Logic

Epistemic logic models knowledge as a modal operator over propositions:

- $K_i\phi$ "Person *i* knows ϕ".
- $K_E\phi$ "The eligible voters all know ϕ" iff *for all e in E:* $K_e\phi$.
- $C_E\phi$ "The eligible voters have common knowledge of ϕ" iff *for all e in E:* $K_E\phi K_E K_E\phi \ldots K_E \ldots K_E\phi$.

Epistemic logic has been extended so that knowledge operators can appear in a first order object language (see, e.g. Baltag et al., 1998).

The argument below makes use of knowledge operators, but is not yet fully formalized, and the specific epistemic logical system to be employed in the final arguments has not yet been defined.

III. Levels of Election Validity

We can now define **basic validity (BV)** for an election as common knowledge among the eligible voters of the following conditions:

- Voter legitimacy **(VL)**:
  - $V \subseteq E$ (voter eligibility, **VE**).
  - *For all reported voters v in V,* $v_a(v) = v$ (voter identity**, VID**).
- Vote integrity **(VIN)**: *For all voters v in V,*
  - $i(v) = s(v)$ (voter competence, **VC**).
  - $s(v) = b_m(v)$ (voting fidelity, **VF**).
  - $b_m(v) = b_c(v)$ (ballot security, **BS**).
  - $b_c(v) = t(v)$ (single tally accuracy, **STA**).
- Result correctness **(RC)**: $R = \sum_{v \in V} t(v)$.

Noncoercible validity (NCV) consists of the validity conditions above together with a new function, a voter's preference $p: V \rightarrow A$, and the following additional vote integrity constraint:

- $C_E\ \forall v \in V: p(v) = i(v)$ (noncoercibility, **NC**).

Finally, **nondeniable validity (NDV)** augments NCV with a two part constraint:

- $C_E \{\forall i\ [K_i\ b_m(v_i) \neq t(v_i)] \Rightarrow C_E\ b_m(v_i) \neq t(v_i)\}$ (nondeniability of the tally, **NDT**); and
- $C_E \{\forall i\ [K_i \sim \textit{Voter legitimacy}] \Rightarrow C_E \sim \textit{Voter legitimacy}\}$ (nondeniability of legitimacy, **NDL**)

IV. Types and Levels of Openness

We can distinguish between two broad types of openness or secrecy in an election: (1) polling openness, related to the identities of voters (*who* votes); and (2) ballot openness, related to the content of the ballots (*for whom or what* each voter votes). Ballot openness can be further broken down into

(a) openness of the marked ballot (b_m);
(b) openness of how the ballot is tallied (function t).

[*Note*: in what follows, I will attach knowledge operators to functional and variable values, rather than propositions. For example, "$C_E\ V$" means "the eligible voters have common knowledge of the set of who has been reported as voting" and "$K_i\ t(v_i)$" means "voter i knows the value of the tally function for his/her own vote."]

We can define the following levels of polling openness:

- Open pollbook **(OP)**: $C_E\ V$.
- Identity visibility **(IV)**:
  $C_E\ \forall v \in V\ K_E\ v_a(v)$.
- Closed pollbook **(CP)**: $C_E \sim \exists i\ K_i\ V$.

We can also define levels of marked ballot openness:

- Self-certifiable **(SC)**: $C_E\ \forall i\ K_i\ b_m(v_i)$.
- Collectively auditable **(CA)**:
  $C_E \sum_{v \in V} b_m(v)$.

As well as levels of ballot tally openness:

- Untraceable **(UT)**: $C_E\ \forall i,j \sim K_i\ t(v_j)$.
- Self-verifiable **(SV)**: $C_E\ \forall i\ K_i\ t(v_i)$.
- Unshareable **(US)**: $C_E\ \forall i,j \neq i \sim K_j\ t(v_i)$.
- Open voting **(OV)**: $C_E\ \forall i,j\ K_i\ t(v_j)$.

V. Sufficient Openness for Different Levels of Election Validity

Proposition 1. *Voter legitimacy (VL) can be satisfied under an open pollbook (OP).*

Argument: Under OP, $C_E\ V$. Therefore all knowledge among the members of E about the membership of E (which is assumed to be collectively sufficient to verify $V \subseteq E$), can be applied, and this is common knowledge, so voter eligibility is satisfied. Each voter can verify whether their membership in V accurately reflects whether or not they voted, and this fact is common knowledge, so that if $\forall v \in V\ v_a(v) = v$, each voter under OP can verify this, and this is common knowledge, so voter identity is satisfied .

Proposition 2. *Basic validity (BV) can be satisfied under open voting (OV).*

Argument: Under OV, $C_E\ \forall i,j\ K_i\ t(v_j)$, viz, each person knows how each voter's ballot was tallied. Therefore, $C_E\ V$, and OP is satisfied, so VL can be satisfied by Propositon 1. Since each voter knows how their own ballot was tallied, if each single tally reflects the voter's intent ($i(v)=t(v)$), then the voters can verify this collectively, and this is common knowledge, so vote integrity (VI) can be satisfied. Because voters all know how each others' ballots were tallied as well, they can all add them up, and if the reported result R (which is assumed to be common knowledge) = $\sum_{v \in V} t(v)$, then the voters can each verify this, and this fact is common knowledge, so result correctness can be satisfied.

Proposition 3. *Noncoercible validity (NCV) can be satisfied if an open pollbook (OP), and self-verifiable (SV) and unshareble (US) ballot tallying all hold.*

Argument: NCV is BV plus noncoercibility (NC): $C_E\ \forall v \in V\colon p(v) = i(v)$. OP satisfies the voter legitimacy (VL) condition of BV by Proposition 1. Under SV, $C_E\ \forall i\ K_i\ t(v_i)$. Therefore each voter can verify $i(v)=t(v)$ for him/herself, and this is common knowledge, so assuming VL holds (meaning that there are no illegitimate votes being tallied), vote integrity can also be satisfied. Assuming that the reported result R is common knowledge, then assuming VL holds leads to common knowledge that R must be accurate if all individuals have indeed verified their own ballot tallies to be individually accurate (VI holds), so result correctness can be satisfied and, with it, BV. US holds by assumption, so $C_E\ \forall i,j{\neq}i\ {\sim}K_j\ t(v_i)$. Since no one in the population can know how someone else voted, there is no way for coercion to work, since a would-be coercer would be unable to tell whether the voter they were trying to coerce voted as instructed. Thus, each voter can intend to vote according to their inner preference, and this is common knowledge, so NC can hold.

Proposition 4. *Noncoercible validity (NCV) can be satisfied if an open pollbook (OP), and self-certifiable (SC) and collectively auditable (CA) ballot marking, and untraceable (UT) ballot tallying all hold.*

Argument: VL can be satisfied under OP through Proposition 1. Under SC, $C_E\ \forall i\ K_i\ b_m(v_i)$. Therefore each voter v can verify for him/herself that $i(v) = b_m(v)$, and this is common knowledge, so voter competence (VC) and voting fidelity (VF) can be satisfied. Under CA, $C_E \sum_{v \in V} b_m(v)$. Therefore, the reported result R can be compared by everyone with $\sum_{v \in V} b_m(v)$, and if they are equal, this outcome is equivalent to establishing ballot security (BS), single tally accuracy (STA), and result correctness (RC). UT corresponds to $C_E\ \forall i,j\ {\sim}K_i\ t(v_j)$, which logically implies $C_E\ \forall i,j{\neq}i\ {\sim}K_j\ t(v_i)$, which is the definition of unshareable (US), and we can therefore apply the same reasoning as in the argument for Proposition 3 to establish that noncoercibility (NC) can be satisfied, thus establishing that NCV can be satisfied as well.

Proposition 5. *Nondeniable validity (NDV) can be satisfied under identity visbility (IV) together with self-certifiable (SC), collectively auditable (CA), and untraceable (UT) ballots.*

Argument: NDV adds the nondeniability conditions to NCV: $C_E\ \forall i\ [K_i\ b_m(v_i) \neq t(v_i)] \Rightarrow C_E\ b_m(v_i) \neq t(v_i)$ (nondeniability of the tally, NDT); and $C_E\ \forall i\ [K_i\ \sim$*Voter legitimacy*$] \Rightarrow C_E \sim$*Voter legitimacy* (nondeniability of legitimacy, NDL). IV means that $C_E\ \forall v \in V\ K_E\ v_a(v)$. This implies OP because if every voter knows the real identity of each member of V, then they must know V as well, and this is common knowledge. Therefore, VL can be satisfied, by Proposition 1. But IV also satisfies NDL, because it implies that any illegitimacy in the membership of V must be common knowledge. SC and CA can satisfy VI and RC by the argument given for Proposition 4. UT satisfies NC by the argument given for Proposition 4. All that remains is to establish NDT. This can be satisfied under UT, because under UT no voter can know $b_m(v_i) \neq t(v_i)$, so the antecedent of the material conditional in the definition of NDT will always be false, therefore NDT will hold.

Proposition 6. *Nondeniable validity (NDV) can be satisfied under a closed pollbook (CP) only if SC, CA, and UT all hold, and if there is a procedure in which voter legitimacy (VL) can be satisfied under CP.*

Argument: SC, CA, and UT can jointly satisfy VI, RC, NC, and NDT by the arguments given for Proposition 5. The definition of CP is $C_E \sim \exists i\, K_i\, V$. This satisfies NDL because no one knows V, and therefore no one can know that VL is violated and so the material conditional in the definition of NDL must hold. But VL must still be satisfied. The proposition simply says that a procedure for this must be found if CP is to satisfy NDV. (If no such procedure is possible, this then becomes an impossibility result.)

VI. Characterizations

Proposition 7. *Nondeniable validity (NDV) can be satisfied if and only if the following conditions hold: self-certifiable (SC), collectively auditable (CA), untraceable (UT), and either (a) identity visibility (IV) or (b) closed pollbook (CP) and voter legitimacy (VL).*

Argument: Sufficiency is established through Propositions 5 and 6. If NDV is satisfied, then VI holds, so $C_E\ \forall v \in V\ i(v) = s(v) = b_m(v)$. Each voter can be assumed to know the value of his/her own intent $i(v)$, and this is common knowledge, so $C_E\ \forall i\ K_i\ b_m(v_i)$ (SC). BS, STA, and RC all hold, so $C_E \sum_{v \in V} b_m(v)$ (CA). NDT holds, so by definition $C_E\ \forall i\ [K_i\ b_m(v_i) \neq t(v_i)] \Rightarrow C_E\ b_m(v_i) \neq t(v_i)$. But since NDV holds by hypothesis, $C_E\ \forall i\ b_m(v_i) = t(v_i)$, so $C_E \sim \exists i\ K_i\ b_m(v_i) \neq t(v_i)$. $C_E\ \forall i\ K_i\ b_m(v_i)$ (every voter is commonly known to know their own ballot mark) because we have established SC, and since no voter knows that this differs from $t(v_i)$, $C_E\ \forall i \sim K_i\ t(v_i)$ (it is common knowledge that no voter knows their own tally value) because if a voter knew this, SV and US would have to hold, viz, the voter would have unique knowledge of their own vote, because otherwise NC would be violated, contra hypothesis. If a voter knew their own vote and could not share it, NDT would not hold, contra hypothesis. Joining $C_E\ \forall i \sim K_i\ t(v_i)$ with $C_E\ \forall i,j \neq i \sim K_j\ t(v_i)$ (US) yields $C_E\ \forall i,j \sim K_i\ t(v_j)$ (UT). NDV implies that either (a) IV or (b) CP with VL must hold because the alternatives are (c) OP but not IV or (d) an intermediate state of polling openness in which neither OP nor CP holds. (c) would violate NDL, because violations of VL could be detected by one voter knowing that V was illegitimate without this being establishable as common knowledge. (d) cannot hold because this would violate VL.

Conjecture 8. *Noncoercible validity (NCV) can be satisfied if and only if either (a) nondeniable validity (NDV) holds, or (b) NDV does not hold but either (i) open pollbook (OP) holds or (ii) CP and VL hold, and either (iii) self-verifiable (SV) and unshareable (US), or (iv) self-certifiable (SC), collectively auditable (CA), and untraceable (UT) hold.*

Conjecture 9. *Basic validity (BV) can be satisfied if and only if either (a) NCV holds or (b) NCV does not hold but OV holds.*

VII. Practical Implications

Implication 1. If coercibility and deniability are not at issue, then open voting works fine. A secret ballot is not necessary, voters can vote outside of controlled settings, at their convenience, and election validity can be established by publishing the list of recorded voters and who/what each voter is tallied as having voted for.

Implication 2. If we are worried about coercibility but not deniability, then

(a) all polling must be controlled (no mail-in or other votes cast outside of a controlled polling place):

(b) the pollbook must be available to everyone (unless we can find a procedure to guarantee voter legitimacy with a closed pollbook) [see the WhoVoted.net website for an example implementation]; and

(c) voters can be allowed to self-verify how their tally was recorded as long as they cannot prove how they voted to anyone else [e.g. by being given a code without an official receipt].

Implication 3. If we are worried about deniability of the tally by one or more voters, then

(a) condition a of Implication 2 must hold;
(b) condition b of Implication 2 must hold;
(c) voters must be allowed to self-certify how their ballot is marked before they submit it [e.g. through a paper ballot];
(d) the result must be auditable by everyone to determine that the marked ballots add up to the reported result [e.g. by keeping paper ballots secured and available for recounting]; and
(e) ballots must be untraceable to individual voters (self-verification cannot be allowed).

Implication 4. If we are worried about deniability of voter legitimacy by one or more voters, then

(a) condition a of Implication 2 must hold;

(b) condition b of Implication 2 must hold; and

(c) if no procedure can be found for guaranteeing voter legitimacy with a closed pollbook, then the identity of each voter must be visible to everyone [e.g. through broadcasting images of each voter entering the voting booth].

Thanks to….

The members of the Voting Technology Working Group of Computer Professionals for Social Responsibility