Heuristic Voting as Ordinal Dominance Strategies

Omer Lev
Ben-Gurion University
Beersheba, Israel
omerlev@bgu.ac.il

Reshef Meir
Technion
Haifa, Israel
reshefm@ie.technion.ac.il

Svetlana Obraztsova
Nanyang Technological University
Singapore
lana@ntu.edu.sg

Maria Polukarov
King's College London
London, United Kingdom
maria.polukarov@kcl.ac.uk

Abstract

Decision making under uncertainty is a key component of many AI settings, and in particular of voting scenarios where strategic agents are trying to reach a joint decision. The common approach to handle uncertainty is by maximizing expected utility, which requires a cardinal utility function as well as detailed probabilistic information. However, often such probabilities are not easy to estimate or apply.

To this end, we present a framework that allows for "shades of gray" of likelihood without probabilities. Specifically, we create a hierarchy of sets of world states based on a prospective poll, with inner sets contain more likely outcomes. This hierarchy of likelihoods allows us to define what we term ordinally-dominated strategies. We use this approach to justify various known voting heuristics as bounded-rational strategies.

Introduction

The question of how agents – human or artificial – choose a strategy when facing a choice, has been at the center of attention in artificial intelligence since its inception. Approaches to decision making often rely on two primary components: the *epistemic state* of the agent (her beliefs on how her actions will affect the world), and her *innate preferences* (the utility or cost associated with each outcome).

In voting scenarios, agents' actions are aggregated to reach a shared result. Voters can make strategic choices once they know what the state of the world is (what other agents are voting), following their own utility function (in most voting settings, an ordinal preference over possible outcomes is assumed). This voting decision may either be applied once based on the current beliefs of the voters, or in an iterative fashion so that voters have several opportunities to observe the world and change their action. When the votes of others are unknown, the epistemic state might depend on some prior knowledge and/or signals from the environment.

The most common way to address this lack of knowledge is to assign probabilities to each state of the world and assume that agents each maximize their expected utility over all possible states (see Related Work). However, in many situations human agents may not have the ability to determine precise probabilities of each state of the world, or to act according to them (Tversky and Kahneman 1974; Chater, Tenenbaum, and Yuille 2006). There is no reason to believe that in voting scenarios people will perform differently in this respect.

Voting without Probabilities Alternative approaches, focusing on decision making in face of *strict uncertainty* (defined in terms of possible or impossible states) have been formulated and applied in various AI and economic settings (Dow and Werlang 1994; Boutilier 1994; Halpern 1997; Matt, Toni, and Vaccari 2009), and more recently, in the context of voting (Conitzer, Walsh, and Xia 2011; Reijngoud and Endriss 2012; Meir, Lev, and Rosenschein 2014). The idea at the core of this approach is that, for any given voter, a vote a' *locally dominates* another vote a if a' is at least as good as a for this voter in any voting profile that she considers to be possible.

In fact, voting behavior can also be defined without any explicit epistemic model. Indeed, several recent papers suggest various heuristics that are specific to a voting rule and/or context. Some of these heuristics have been shown to be empirically consistent with voters' behavior in lab experiments (Laslier 2010), and others guarantee desirable convergence and/or welfare properties when applied by all voters in a group (Grandi et al. 2013).

Contribution We extend the framework suggested in the strict uncertainty papers mentioned above, by allowing **gradual levels of uncertainty**. Specifically, we build on the idea of having a likelihood hierarchy – a sequence of sets of states of the world, where each next set is a superset of the previous set in the sequence, so that the states in inner sets are considered by the voter to be more likely. An *undominated vote* in this setting is one which is not dominated at any level of the hierarchy.

Using this hierarchy of likelihood, we suggest **an alternative representation for information structures in voting**. We show how the relevant information can be boiled down to what we call a *pivot-graph*, which succinctly captures all situations where the voter may be pivotal. We then show that the information structure allows us to justify several existing voting heuristics as rational decisions for an appropriate epistemic model (a specific hierarchy of pivot-graphs).

This observation enables us to generalize existing convergence results in the literature on iterative voting, by showing how convergence follows from topological properties of the pivot-graphs.

Related Work For an up-to-date coverage of iterative voting, heuristics and uncertainty-based models, see (Meir 2017). In particular, Conitzer et al. (2011) consider a voter facing an arbitrary information set, and Reijngoud and Endriss (2012) study partial information settings where, for example, only the candidates' scores or only the identity of the leader are known. Closest to our paper is the *local dominance* model (Meir, Lev, and Rosenschein 2014), in which all voters base their belief on a shared *prospective state*. It has been shown that in an iterative voting setting where voters play possible actions that dominate their current action, they are guaranteed to converge to an equilibrium under certain assumptions on the distance metric (Meir, Lev, and Rosenschein 2014; Meir 2015).

Voting heuristics do not explicitly define voters' beliefs; instead, they specify a (typically) simple function that dictates a vote in every given state, aiming to capture realistic voting behaviors (Reijngoud and Endriss 2012; Grandi et al. 2013). In particular, some models suggest that a non-pivotal voter either votes truthfully (Dutta and Laslier 2010) or abstains (Desmedt and Elkind 2010).

These models stand in contrast with the expected utility models, such as, for example, the *calculus of voting* (Riker and Ordeshook 1968; Myerson and Weber 1993), where a voter computes the probability for each action (vote) to be pivotal in every pairwise tie. We see our model as a way to capture a similar line of reasoning in identifying the influential ties, albeit without using probabilities. A more fundamental difference with the calculus of voting approach is that the latter assumes a common knowledge of rationality and the preference distribution, from which an equilibrium is derived.

In the more general AI literature, there are other non-probabilistic models of uncertainty, where two of the most prominent ones are the *possibility theory* (Dubois, Fargier, and Prade 1996) and *Dempster-Shafer theory* (Shafer 1976). These models attribute a cardinal possibility measure to states and develop calculus rules for belief updates and comparisons. The closest to our work is the *plausibility measure* approach (Halpern 1997), that allows for a partial order of plausibility. Our hierarchical ordinal dominance concept is even more strict, and relies on the structure of the problem where uncertainty is essentially about the accuracy of a single point estimate (a poll).

While we take our distance-based epistemic assumptions from the aforementioned local dominance voting model, an earlier precursor of this idea is the logic for inexact knowledge based on *margin-of-error* (Williamson 1992).

Model

An election is composed of a set V of n voters and a set C of m candidates. Each voter $i \in V$ has a weak preference relation $\succsim^i \subseteq C \times C$ over the candidates, that is, for each two candidates $x, y \in C$, $x \succsim_i y$ or $y \succsim_i x$, and if both are true, they are equivalent. Moreover, the relation is transitive (so, for $x, y, z, \in C$, $x \succsim_i y$ and $y \succsim_i z$, then $x \succsim_i z$).

The voting rules we shall focus on are the *score-based voting* (SBV) rules. An SBV $(\hat{f}, A)$ is defined by a set $A \subseteq \mathbb{N}^m$ of allowed votes, and a function $\hat{f} : A^n \to C$. For example, the set A under Plurality contains all vectors that have exactly one non-zero element, which is 1; Approval allows all binary vectors; Borda allows all permutations of $(0, \dots, m-1)$; etc. We denote by $\vec{a} = (a_i)_{i \in V}$ the *voting profile*; by $a_i(c) \in \mathbb{N}$ the absolute score given to c by agent i in vote $a_i \in A$; and by $s_{\vec{a}}(c) = \sum_{i \in V} a_i(c)$ the total score given to candidate c. The winner is $\hat{f}(s_{\vec{a}}) = \arg\max_{c \in C} s_{\vec{a}}(c)$, breaking ties lexicographically.

For each voter i, the outcome (and thus, her utility) depends on her own vote, as well as on the state of the world $s = s_{\vec{a}_{-i}}$, that encompasses the votes of all other participants. We separate these two arguments by writing the outcome function as $f(s, a_i) = \hat{f}(s + a_i)$. For every state s and any two actions $a_i, b_i \in A$, we write $f(s, a_i) \succ^i f(s, b_i)$ when voter i prefers action a_i over b_i at state s.

Example 1. *There are 100 voters and 5 candidates – w, b, c, d, e – using the Plurality voting system. The voters have access to a poll where votes are $\overline{s} = (29, 26, 22, 17, 6)$. A voter i currently voting for b sees the state $s = s_{\vec{a}_{-i}} = (29, \mathbf{25}, 22, 17, 6)$. For any action a'_i of i, $f(s, a'_i) = w$, which means that voter i is indifferent between her actions.*

Note that the voters never explicitly reason about the preferences of other individuals – only about their (aggregated) actions. We will return to this example later in the paper.

Information Structures

An *information set* is a set of states $S' \subseteq S$. An *information structure* of agent i is a collection of information sets $\mathcal{S}^i = (S^i_j)_{j=1}^k$, where $S^i_j \subseteq S^i_{j+1}$ for all j. That is, each information set contains the sets with a lower index.

An agent does not assign probabilities to states or to information sets, but an intuitive interpretation of the model is that agent i believes any state in S^i_j to be *substantially more likely* than all states outside S^i_j. An information structure can either be shared by all agents, or be agent-specific.

Example 2. *Consider voter i from Example 1 and assume she has an information structure $\mathcal{S}^i = (S^i_1, S^i_2)$ as in Figure 1. In particular, i believes that candidate c may win (as, e.g., $s' = (24, 21, 26, 17, 6) \in S^i_2$); however, this is far less likely than a victory of w or b, as there is no state $s'' \in S^i_1$ where c wins.*

Ordinal Dominance

Following (Conitzer, Walsh, and Xia 2011; Reijngoud and Endriss 2012; Meir, Lev, and Rosenschein 2014), for any information set S^i_j and actions $a, b \in A$, we say that action a S^i_j*-dominates* action b (denoted $a \succ^i_j b$) if $f(s, a) \succsim^i f(s, b)$ for *all* $s \in S^i_j$ and $f(s, a) \succ^i f(s, b)$ for *at least one* $s \in S^i_j$. Agent i is *indifferent* between actions a, b at S^i_j (denoted $a \sim^i_j b$) if $f(s, a) \sim^i f(s, b)$ for all $s \in S^i_j$.

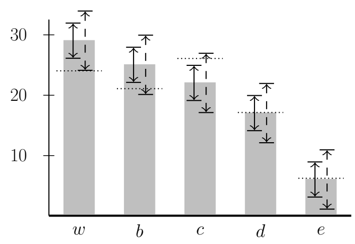

Figure 1: Information structure of voter i at state $s_{\vec{a}_{-i}} = (29, \mathbf{25}, 22, 17, 6)$ (gray pillars). The set S_1^i contains all states that result from changing the score of any candidate in $s_{\vec{a}_{-i}}$ by at most 3 votes (solid arrows). The set S_2^i allows for a variation of 5 votes (dashed arrows). The dotted lines indicate state s' from Example 2, where c wins.

Note that S_j^i-dominance is a *partial order* over actions A (transitive, antisymmetric and irreflexive relation).

Definition 1. *Action* a ordinally dominates *action* b *(in structure* $\mathcal{S}^i = (S_j^i)_{j \in [k]}$*) if there is some* $j \in [k]$ *such that action* a S_j^i*-dominates action* b*.*

The next lemma guarantees that it is not possible that $a \succ_j b$ and $b \succ_{j'} a$ for some $j' \neq j$.

Lemma 3. *Ordinal dominance is a partial order.*

Proof. Transitivity: Suppose action a ordinally dominates action b and b ordinally dominates action c, due to S_j^i and $S_{j'}^i$, respectively. W.l.o.g. $j' \leq j$, then $S_{j'}^i \subseteq S_j^i$. There is a state $s' \in S_{j'}^i$ where $f(s', b) \succ^i f(s', c)$, and since $s' \in S_{j'}^i \subseteq S_j^i$, we also have $f(s', a) \succeq^i f(s', b)$, and so $f(s', a) \succ^i f(s', c)$. Similarly, for any $s \in S_{j'}^i$, $f(s, a) \succeq^i f(s, b) \succeq^i f(s, c)$. Thus $a \succ_{j'}^i c$ which means that a ordinally dominates c.

Antisymmetry: Suppose action a ordinally dominates action b due to S_j^i. For every $j' \leq j$, there cannot be a state $s \in S_{j'}^i \subseteq S_j^i$ where $f(s, b) \succ^i f(s, a)$. Similarly, for any $j' > j$, there is a state $s' \in S_j^i \subseteq S_{j'}^i$ where $f(s', b) \prec^i f(s', a)$. Thus, b does not $S_{j'}^i$-dominates a. Since this is true for any j', b does not ordinally dominate a. □

Distance-Based Uncertainty

Following Meir et al. (2014), we consider the following way to derive information sets and information structures. Given a metric $d : S \times S \to [0, 1]$ and a parameter $r \in [0, 1]$, every state $s \in S$ explicitly defines an information set $S_{d,r}(s) = \{s' : d(s, s') \leq r\}$. In general, the metric d can be completely arbitrary and the induced set is meaningless.[1] However, in the context of voting there are several natural metrics: For example, $d(s, s')$ may reflect what fraction of votes has changed between s and s'. In Meir et al. (2014) and Meir (2015), the distances between candidate score vectors were defined by different ℓ-norms and the Earth Mover distance (EMD), which is essentially the ℓ_1 norm with the additional constraint that the total number of votes remains the same. Thus, $S_{d,r}(s)$ may reflect a range of possible candidates' scores given a poll or a current state s.

[1]In fact, *any* set $S' \subseteq S$ can be derived from s for some carefully designed metric d.

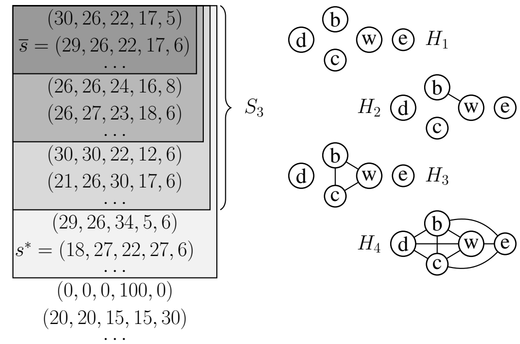

Figure 2: A schematic example of the information structure $\mathcal{S}_{EMD,\vec{r}}(\overline{s})$, and the induced pivot graph structure $\mathcal{H}_{EMD,\vec{r}}(\overline{s})$. E.g., graph H_4 contains the edge (b, d) due to state $s^* \in S_4$. States below S_4 are considered impossible. All the pivot graphs are upward closed w.r.t. the lexicographic order on C, but they are not always a clique ($(d, e) \notin H_4$).

Pivot Graphs

A pair of actions (a', a'') is *pivotal* for a pair of candidates $c', c'' \in C$ in state $s \in S$, if $f(s, a') = c'$ and $f(s, a'') = c''$. An agent i is pivotal for the pair of candidates $c', c'' \in C$ in information set S_j^i, if there are $s \in S_j^i$ and actions $a_i', a_i'' \in A$ that are pivotal for c', c'' in s.

Information set S_j^i then induces a *pivot-graph* $H_j^i = (C, E)$, which contains a vertex for every candidate, and an edge (c', c'') if agent i is pivotal for the pair c', c'' in S_j^i.

Every information structure $\mathcal{S}^i$ induces a *pivot graph structure* $\mathcal{H}^i = (H_j^i)_{j=1}^k$, where each H_j^i is a subgraph of H_{j+1}^i (since adding more states can only add edges to the graph). The set $\mathscr{H}(C)$ contains all pivot graph structures.

Epistemic Models and OD Equilibrium

An *epistemic model* of agent i maps any state s to an information structure $\mathcal{S}^i(s) = (S_1^i(s), S_2^i(s), \ldots, S_k^i(s))$, and thus also to a pivot graph structure $\mathcal{H}^i(s)$.

We define the set $OD_{\succsim^i}(\mathcal{S}, a)$ that contains all actions that ordinally dominate a in $\mathcal{S}$ according to preferences $\succsim^i$, and a set $UOD(\mathcal{S}, a)$ containing all actions a' that ordinally dominate a and are not ordinally dominated themselves. Naturally, this leads to a definition of an OD-equilibrium – when for every agent $OD_{\succsim^i}(\mathcal{S}^i(s_{\vec{a}_{-i}}), a_i)$ is empty (and hence, $UOD_{\succsim^i}(\mathcal{S}^i(s_{\vec{a}_{-i}}), a_i) = \emptyset$, too).

Observation 4. *For a "full information" epistemic model where* $\mathcal{S}^i(s) = (\{s\})$*, the set* $OD(\mathcal{S}^i(s), a_i)$ *coincides with the set of better-responses to* (s, a_i)*; the set* $UOD(\mathcal{S}^i(s), a_i)$ *coincides with best-responses; and OD equilibrium coincides with a pure strategy Nash equilibrium.*

An epistemic model is *cliqued* if its mapping to a pivot graph structure $\mathcal{H}(s)$ at every state $s \in S$, $H_j(s)$, is a clique. The epistemic model is *upward closed* if the pivot graph structure $\mathcal{H}(s)$ at every state s has an order L over candidates such that if $(c, c') \in H_j(s)$ and $c'' >_L c'$ then $(c, c'') \in H_j(s)$. Note that any cliqued epistemic model is upward closed (where L may be an arbitrary order where all candidates in $H_j(s)$ precede all others). More generally, L can be roughly thought of as an order of likelihood of states. For simplicity of notation, we denote both the pivot graph structure and the epistemic model (which is a function from states to structures) by $\mathcal{H}$.

Distance metrics provide us with a simple way to define an information structure: given a metric d and an increasing sequence of distances $\vec{r} = (r_1, r_2, \dots, r_k)$, we get an epistemic model $\mathcal{H}_{d,\vec{r}}(s) = (H_{d,r_1}(s), H_{d,r_2}(s), \dots, H_{d,r_k}(s))$ that is induced by $\mathcal{S}_{d,\vec{r}}(s)$. We later show how the topological properties of $\mathcal{H}_{d,\vec{r}}(s)$ are related to the properties of the metric d.

Example 5. *We expand Example 1 where candidates' scores (in % of total) are* $(29, 26, 22, 17, 6)$. *We do not specify the number of voters in the poll. We consider a voter with a concentric information structure, based on the radii* $\vec{r} = \{1\%, 3\%, 7\%, 17\%\}$ *and the EMD metric. These information sets induce pivot-graphs as illustrated in Figure 2.*

In Example 5, consider a Plurality voter whose preferences are $e \succ d \succ c \succ b \succ w$. Then the action "$c$" (a shorthand for $(0, 0, 1, 0, 0)$) ordinally dominates action "e" (due to H_3) and "b" ordinally dominates everything else due to H_2.

Sharp Pivot Property in Large Populations

Note that structures $\mathcal{S}^i$ and $\mathcal{H}^i$ are two different ways to represent the information of an agent. In general, $\mathcal{H}^i$ contains less information than $\mathcal{S}^i$, since there may be states where some action pairs are pivotal, whereas others are not.

Yet, it seems plausible that in a large population, a voter is unlikely to make such fine distinctions about the possible outcomes: voters do not know the exact score of each candidate, but only have a rough idea of what it is (each candidate's share of the votes). As a result, it is reasonable to assume that if a voter considers herself pivotal in some possible tie, she will consider *any* change in her vote as possibly pivotal. We capture this intuition in the following formal definition.

Definition 2 (Sharp Pivot Property (SPP))**.** *An information structure $\mathcal{S}^i$ satisfies the* Sharp Pivot Property *if: for all j and all $c', c'' \in C$, an edge $(c', c'') \in H^i_j$ entails that any pair of actions (a'_i, a''_i) such that $a'_i(c') - a'_i(c'') > a''_i(c') - a''_i(c'')$ is pivotal for (c', c'') in S^i_j.*

That is, if there is some action pair that makes c' the winner instead of c'', then *any move* that increases the gap in favor of c' might make c' beat c'' and become the winner (in some $s \in S^i_j$).

Working with pivot-graph structures is much more convenient than working with arbitrary information sets, and their meaning in the context of voting is clear. We will assume throughout the paper that all information structures have SPP, which means that $\mathcal{H}^i$ contains all the relevant information in $\mathcal{S}^i$.

Justifying SPP While SPP is plausible, we would like to show that at least in some cases it provably holds. The argument is a rather fine one, that relies on viewing the (finite) poll as approximating some underlying real-valued distribution $p \in \Delta(C)$. Distribution p defines a unique score vector s^n_p for every population size n (using some fixed rounding of $p \cdot n$), and we argue that for a sufficiently large population n, the information structure $\mathcal{S}_{d,\vec{r}}(s^n_p)$ satisfies SPP.

Intuitively, consider the leader w in Example 5 with the Plurality rule, some other candidate (say b), and some information set $S_{d,r_j}(s)$. If $s(w) - s(b)$ is much lower than $r_j \cdot n$, then any move where the voter deserts w and/or joins b might be pivotal for (b, w). In contrast, if $s(w) - s(b)$ is much higher than $r_j \cdot n$, then b can never win or even tie with w. Only if $s(w) - s(b) \cong r_j \cdot n$, then moving from w to b makes b the winner, but (say) moving from d to b does not. This extreme case becomes unlikely as the population gets larger.

Theorem 6. *For the EMD metric d, any SBV, and any radii vector $\vec{r}$, the following holds for almost all[2] distributions $p \in \Delta(C)$: There is n_0 such that for all $n > n_0$, the information structure $\mathcal{S}_{d,\vec{r}}(s^n_p)$ satisfies SPP.*

Computing Dominance Relations

We show that strategies can be efficiently compared according to ordinal dominance.

Proposition 7. *Given a pivot graph structure $\mathcal{H}^i = (H^i_1, \dots, H^i_k)$ and any SBV $(\hat{f}, A)$, voter i can check in time $O(m^2 k)$ if vote $a'_i \in A$ ordinally dominates vote $a_i \in A$.*

Intuitively, Algorithm 1 checks (for each uncertainty level j), whether the new vote a'_i is "safe" (not worse than a_i in any possible tie), and whether it is "pivotal" (better than a_i in at least one tie).

$I[X] \in \{-1, 1\}$ is an indicator variable for statement X, and we use $a_i(c)$ to indicate candidate c's score when voter i vote is a_i. The complexity of checking whether a given vote a_i is *undominated* is left as an open question.

Algorithm 1: $\text{OD}(a'_i, a_i \in A, \succsim^i, \mathcal{H}^i \in \mathscr{H})$

for $c, c' \in C$ **do**
- $\text{diff}(c, c') \leftarrow a'_i(c) + a_i(c') - a_i(c) - a'_i(c')$;
- $\text{effect}(c, c') \leftarrow \text{sign}(\text{diff}(c, c') \cdot I[c \succ^i c'])$;

for $j \leq k$ **do**
- $\text{safe}(j) \leftarrow \min_{(c,c') \in H^i_j} \text{effect}(c, c')$;
- $\text{pivot}(j) \leftarrow \max_{(c,c') \in H^i_j} \text{effect}(c, c')$;
- $\text{dom}(j) \leftarrow I[\text{pivot}(j) + \text{safe}(j) \geq 1]$;

if $\exists j \leq k$ *s.t. dom*$(j) = 1$ **then**
- **return** TRUE

else
- **return** FALSE

[2] Except a 0-measure set of distributions.

Proof. Suppose a' ordinally dominates a. Then there is some level $j \leq k$ such that $a' \succ_j^i a$. This means that for any pair of candidates (c, c') that can be tied in H_j^i, either c is preferred to c' and a' weakly reduced c''s score, or c' is preferred to c and a' weakly adds to c''s score (thus effect$(c, c') \geq 0$). Hence, in particular safe$(j) \geq 0$. In addition, there must be a pair for which the gain is strict, and effect$(c, c') = 1$, which means pivot$(j) = 1$. In total, dom$(j) \geq 1 + 0 = 1$ so the algorithm returns TRUE.

Otherwise, in every level j, either a'_i, a_i have the same outcome in all states, or there is a pair $(c, c') \in H_j^i$ such that $f(s, a_i) = c, f(s, a'_i) = c'$, and $c \succ^i c'$.

In the latter case, since f is a scoring rule this means that $a'_i(c) - a_i(c) < a'_i(c') - a_i(c')$, i.e. that c' gained strictly more score than c when changing from a_i to a'_i. Thus diff$(c, c') = a'_i(c) + a_i(c') - a_i(c) - a'_i(c') < 0$, and effect$(c, c') = -1$. The algorithm then computes safe$(j) = -1$. Therefore dom$(j) \leq 1 - 1 = 0$.

In the first case, effect$(c, c') = 0$ for all pairs, and thus safe$(j) =$ pivot$(j) = 0$, and dom$(j) = 0$. □

Justifying Voting Heuristics with OD

Many heuristics have been suggested to analyze how voters behave and change their vote. Most heuristics are derived from a single "prospective state" s, which is assumed to be the current voting profile or poll. Formally, a *set heuristics* is a function $h : S \times A \to 2^A$ that maps the prospective state and the current action to a set of new possible actions. We say that h is a *point heuristics* if $|h(s, a)| \leq 1$ for every s, a. To be consistent with previous definitions, we always omit a from the set $h(s, a)$, and assume that when $h(s, a) = \emptyset$ the voter simply keeps her current vote.

Definition 3. *We say that an epistemic model $\mathcal{H}$* justifies *heuristic h, if for any state $s \in S$ and current action $a \in A$: (I) $h(s, a) = \emptyset$ if and only if $UOD(\mathcal{S}(s), a) = \emptyset$; and (II) $h(s, a) \subseteq UOD(\mathcal{S}(s), a)$. $\mathcal{H}$* strongly justifies *h if (II) holds with equality.*

This means that the heuristic h always recommends undominated ordinal-dominance moves under the epistemic model $\mathcal{H}$, and only keeps the current action if no such move exists.

As a simple example, consider the Plurality rule and the heuristic $h^{not-last}(s, a)$ that is empty except when action a is the least preferred candidate $\hat{a}_i$, and then it moves to an arbitrary other candidate. Consider the epistemic model $\mathcal{H}^{all}(s) = (H_1^{all})$ where H_1^{all} is the complete graph.

Observation 8. *$\mathcal{H}^{all}$ strongly justifies $h^{not-last}$.*

This is since (I) suppose that $a \neq \hat{a}_i$. Then no candidate ordinally dominates a and thus $UOD(\mathcal{H}^{all}(s), a) = \emptyset = h^{not-last}(s, a)$; (II) when $a = \hat{a}_i$, any other candidate c is undominated but globally dominates a (since there is a possible state where i is pivotal for c against a), in which case $UOD(\mathcal{H}^{all}(s), a) = C \setminus \{\hat{a}_i\} = h^{not-last}(s, a)$.

Local Dominance

Local dominance (Meir, Lev, and Rosenschein 2014) heuristic with metric d and parameter r explicitly define a set $S_{d,r}(s) = \{s' : d(s, s') \leq r\}$. The heuristic action $h_{d,r}^{LD}(s, a_i)$ is defined for the Plurality rule as follows: Let $D \subseteq C$ be the set of candidates that $S_{d,r}(s)$-dominate a_i; If D is non-empty, then vote for the most preferred candidate in D.

We define an epistemic model $\mathcal{H}_{d,r}^{LD}$ where $\mathcal{H}_{d,r}^{LD}(s)$ contains a single pivot graph H_1 which is the pivot graph induced by $S_{d,r}(s)$. Note that our definition applies for *any voting rule*, unlike the one in Meir et al. (2014). In Plurality, $\mathcal{H}_{d,r}^{LD}$ justifies $h_{d,r}^{LD}$ (straightforward proof omitted due to space constraints).

Truth/lazy-bias Denote the top candidate of i by $q_i \in C$, and denote by $\perp$ an "abstain" action that adds no score to candidates. We adopt the suggested variations in Dutta and Laslier (2010) and Desmedt and Elkind (2010), where the voter prefers the truthful/abstain action if this does not affect the outcome. However, this naïve modification alone may lead to unreasonable behaviors, e.g., where no-one votes (Elkind et al. 2015), even under full information.[3]

For $r_2 > r_1$, the "truth bias" heuristics $h_{d,r_1,r_2}^{LD+TB}(s, a_i)$ is as follows (Meir, Lev, and Rosenschein 2014): (1) agent i performs a local-dominance move at radius r_1, if exists. If such a move does not exist, i examines if $f(s', a_i) \succ^i f(s', q_i)$ for some $s' \in S_{d,r_2}(s)$. (2a) If so, agent i keeps the current vote a_i, (2b) otherwise, i moves to q_i.

While the behavior seems to maintain the reason behind truth bias, the definition of h is cumbersome. Instead, we can use r_1, r_2 to define an epistemic model $\mathcal{H}_{d,r_1,r_2}^{LD+TB}$ as follows. We let $H_1(s)$ be as in $\mathcal{H}_{d,r_1}^{LD}$ above. We similarly compute $H_2(s)$ from $S_{d,r_2}(s)$, but taking only edges between the current vote a_i and candidates less preferred than a_i. Let $\mathcal{H} = \mathcal{H}_{d,r_1,r_2}^{LD+TB}(s) = (H_1(s), H_1(s) \cup H_2(s))$.

Proposition 9. *$\mathcal{H}_{d,r_1,r_2}^{LD+TB}$ justifies h_{d,r_1,r_2}^{LD+TB} in Plurality.*

Proof. First, if $H_1(s)$ is nonempty (at least one tie) then $h_{d,r_1,r_2}^{LD+TB}(s, a_i) = a_i^* \in UOD(\mathcal{H}(s), a_i)$ as in a standard LD move. Otherwise, there are two cases.

If $H_2(s)$ contains some edge (a_i, b), then by SPP for any $a' \neq a_i$ there is a state $s' \in S_{d,r_2}(s)$ where $f(s', a_i) = a_i$ and $f(s', a') = b$ (think of s' as state where a single additional vote for a_i is critical). Since a_i is preferred to b by the definition of $H_2(s)$, we conclude that no candidate a' dominates a_i in $H_2(s)$ (thus $UOD(\mathcal{H}(s), a_i) = \emptyset$); and that $f(s', a_i) = a_i \succ^i b = f(s', q_i)$ (thus $h_{d,r_1,r_2}^{LD+TB}(s, a_i) = \emptyset$).

In the second case, there is no such edge, then $H_2(s)$ is empty. This means that no action of i can change the outcome whatsoever, and thus by the slight truth-bias q_i is strictly preferred to any other action. In particular, it ordinally dominates a_i and is undominated so $UOD(\mathcal{H}(s), a_i) = \{q_i\}$. Finally, since i is non-pivotal then in particular there is no state $s' \in S_{d,r_2}(s)$ such that $f(s', a_i) \succ^i f(s', q_i)$. Thus $h_{d,r_1,r_2}^{LD+TB}(s, a_i) = q_i \in UOD(\mathcal{H}(s), a_i)$, as required. □

[3] To be completely formal, the preference relation $\succsim^i$ has to be extended to preferences over pairs (winner,action).

The statement for lazy-bias is similar, and uses the same information structure but with a slight preference to abstain instead of voting truthfully.

T-Pragmatist

The T-pragmatist (point) heuristic (Brams and Fishburn 1978; Reijngoud and Endriss 2012) considers the leading T candidates in s (denoted $\mathbf{T}$), and sets a new action $a'_i = h^{T-prag}(s, a_i)$ where a'_i is identical to a_i except the favorite candidate in $\mathbf{T}$ is given maximal score (in any SBV). E.g., in Example 1, if $e \succ^i c \succ^i b \succ^i w \succ^i d$, then $h^{2-prag}(s,a_i)=b$ and $h^{3-prag}(s,a_i)=h^{4-prag}(s,a_i)=c$.

Consider a single-level epistemic model i $\mathcal{H}^{T,i-star}(s)$, which contains a star graph, in which the center node is the most preferred candidate by voter i in the top T candidates, and it is tied with all other $T-1$ candidates in the top T.

It is possible to show that (I) $h^{T-prag}(s,a_i) = \emptyset \iff OD(\mathcal{H}^{T,i-star}(s), a_i) = \emptyset$; and (II) $h^{T-prag}(s,a_i) \subseteq OD(\mathcal{H}^{T,i-star}(s), a_i)$. This shows a connection between the heuristic and the epistemic model, but it is not a sufficient justification since $h^{T-prag}(s,a_i)$ may be dominated. A closer look reveals that the actions dominating it are quite plausible: ranking the other candidates in $\mathbf{T}$ at the bottom can only benefit the voter! We conclude that the T-pragmatist heuristic could be improved.

We define the h^{T*} heuristic similarly to h^{T-prag}, with the following difference: all other candidates in $\mathbf{T}$ *get minimal score* (i.e., ranked at the bottom of a'_i) while maintaining the same order among themselves as in a_i.

Proposition 10. *$\mathcal{H}^{T,i-star}$ strongly justifies h^{T*} in any SBV.*

Proof. Denote by q'_i the favorite candidate of i in $\mathbf{T}$. For $T=1$, the graph has no edges, which means the heuristic advises doing nothing as well. Otherwise, the graph contains all $T-1$ edges (q'_i, x) for $x \in \mathbf{T} \setminus \{q'_i\}$.

Let $a'_i = h^{T*}(s,a_i)$ be any vote s.t. $a'_i(q'_i) \geq a'_i(y) \geq a'_i(x)$ for all $x \in \mathbf{T}\setminus\{q'_i\}$, $y \notin \mathbf{T}$, and $a_i(x) > a_i(x') \Rightarrow a'_i(x) > a_i(x)$.

- If a_i has the same inequalities then $h^{T*}(\vec{s}, a_i) = \emptyset$.
- We have that for all $x \in \mathbf{T}\setminus\{q'_i\}$, $a'_i(q'_i) - a'_i(x) > a_i(q'_i) - a_i(x)$ so in a tie (q'_i, x) vote a_i never beats a'_i.
- If a_i does not rank q'_i at the top, then consider x' ranked lowest by a'_i. We have that $a'_i(q'_i) - a'_i(x') > a_i(q'_i) - a_i(x')$, thus there is a near-tie state $\vec{s}'$ of (q'_i, x') where $f(\vec{s}', a'_i) = q'_i \succ^i x' = f(\vec{s}', a_i)$.
- Similarly, if a_i does not rank some $x' \in \mathbf{T}\setminus\{q'_i\}$ at the bottom, then in particular $a'_i(x') < a_i(x')$ and thus in a near-tie of (q'_i, x'), vote a'_i beats a_i.

Finally, let b be any other vote. We need to show that a'_i is not ordinally dominated by b. b can only beat a'_i in some $s' \in H$ if there is a possible (near-)tie (q'_i, x), and $b(x) < a'_i(x)$. Consider x' that is ranked at the bottom by a'_i. It must be that $b(x') > a'_i(x)$. Thus in a near-tie of (a'_i, x'), vote a'_i beats b. □

Leader Rule (Approval voting)

Assume candidates $c_1, \ldots, c_m$ are sorted in decreasing score order in a state s. In Approval voting the allowed actions are $A = 2^C$. The *Leader rule* (Laslier 2009) $a' = h^{LR}(s, a_i)$ is a strategy approving all candidates strictly preferred to the leader of s, and approves the leader of s (candidate c_1) if and only if it is preferred to the runner-up c_2 (i.e., exactly one of c_1, c_2 is being approved in a').

We consider the epistemic model where $\mathcal{H}^{LR}(s)$ consists of two nested pivot graphs. The inner graph H_1 contains a single edge between c_1 and c_2. The outer graph H_2 is a star connecting c_1 to all candidates.

Proposition 11. *$a' = h^{LR}(s,a_i)$ ordinally dominates all other actions according to $\mathcal{H}^{LR}$. In particular, $\mathcal{H}^{LR}$ strongly justifies h^{LR}.*

Proof. Let a'' be any alternative vote to a'. We will show that a' dominates a'' in at least one of the tie graphs H_1 or H_2.

Consider a'' that differs from a' on (at least) c_1 or c_2 or both. On the graph H_1, the voter is pivotal for c_1, c_2 and thus there is a state s where $f(s,a'') = c_2 \prec^i c_1 = f(s,a')$, or $f(s,a'') = c_1 \prec^i c_2 = f(s,a')$. Thus a' dominates a'' on H_1.

Next, consider a'' that approves c_1, c_2 iff a' approves them, but differs in (at least) some other candidate c'. If $c_1 \succ c'$, c' is not approved in a' and thus approved in a'' (this is regardless of whether c_1 is approved). Since there is a state s in H_2 where c_1 and c' are tied, $f(s,a'') = c' \prec c_1 = f(s,a')$. If $c_1 \prec c'$, c' is approved in a' but not in a''. Again, since there is a state s where they are tied, $f(s,a') = c' \prec c_1 = f(s,a'')$. Thus $a' \succ^i_2 a''$ and therefore a' ordinally dominates a''. □

OD and Iterative Voting

Since ordinal-dominance induces a natural concept of OD-response, we are interested in its implications on iterative voting with multiple strategic voters. In iterative voting, voters proceed from some initial state s^0, and in each iteration an arbitrary voter changes her vote, a process that may either converge to an equilibrium or reach a cycle. Our convergence results depend on the structure of the pivot graphs in the epistemic model.

We first show that both cliqued and upward-closed epistemic structures are the result of distance-based uncertainty with natural assumptions on the distance function.

Proposition 12. *1. Any neutral distance metric d on scoring vectors induces an upward-closed epistemic model.*

2. Any candidate-wise distance metric[4] d on scoring vectors induces a cliqued epistemic model.

Proof of 1. Assume for contradiction that there is a state s in which there are $c_1, c_2, c_3 \in C$ such that $s(c_1) \geq s(c_2) \geq s(c_3)$, c_2, c_3 are tied in a state s' within a distance r from s, but c_1, c_3 are not tied within the same distance. Let $\vec{w}$ be the

[4] This is a metric on a scoring vector, composed of a singleton metric $D : \mathbb{R}^2 \to [0,1]$, where $d(s,s') = \max_{c\in C} D(s(c), s'(c))$. This includes, for example, the ℓ_∞ norm.

score vector for s, and let $\vec{w'}$ be scoring vector for s'. From the triangle inequality, $d(\vec{w}-\vec{w'},0) = d((c'_1,\ldots,c'_m),0) \leq r$. We now examine the $\tilde{\vec{w}}(c'_2+w_1-w_2,0,c'_3,c'_4,\ldots,c'_m)$. The first element has to be smaller than c'_1, and since $\sum_{i=1}^{m} c'_i = 0$, we now begin reducing $c'_4,\ldots,c'_m$ until we create a= $\tilde{\boldsymbol{w}}'$, such that its elements sum up to 0 as well. Since every dimension in the new vector is less than before, $d(\tilde{\boldsymbol{w}}',0) \leq d(\vec{w}-\vec{w'},0) \leq r$, and $\tilde{\boldsymbol{w}}' + \vec{w}$'s has a tie between c_1 and c_3. Proof of 1: Assume that there is a state $s=(s_1,\ldots,s_m)$ in which there are $c_1,c_2,c_3 \in C$ such that $s_1 \geq s_2 \geq s_3$; and another state s' within a distance r from s where c_2,c_3 are tied. We construct a (non-normalized) vector s'' where c_1,c_3 are tied, such that $|s''_j - s_j| \leq |s'_j - s_j|$ for all j (hence s'' is closer to s than s') or one where s'' is such that $s''_j = s'_j$ for $j>2$ and $|s''_1 - s_1| \leq |s'_2 - s_2|$ and $|s''_2 - s_2| \leq |s'_1 - s_1|$.

W.l.o.g we have $s'_3 \geq s_3$ and $s'_j \leq s_j$ for all $j \neq 2,3,$. Denote by $w = s'_2 = s'_3$ be the winning score in s'. There are several cases: (I) if $w \geq s_1$, define $s''_1 = s_1, s''_2 = s_2 \leq s''_1, s''_3 = s_1$; (II) if $s_2 < w < s_1$, define $s''_1 = w \in [s'_1+1, s_1], s''_2 = s_2, s''_3 = w$. It is easy to check that s'' holds both conditions, thus $d(\frac{s''}{\|s''\|},s) \leq d(s',s) \leq r$ as required. If (III) $w \leq s_2$, it quite simple to see that by setting $s''_1 = w, s''_2 = w-1, s''_3 = w$ we are closer to s than s'.

Proof of 2: any two candidates which are tied with the score leader of s – c_1 – at states at distance r from s are also tied for the leadership in a state within the same distance r from s. Since if there is any tie between candidates c', c'', either one of them is c_1, or both of them are tied with c_1 in the radius r (as the difference in the score of c_1 in s and the state where it isn't tied for the victory is larger than when it is tied for the win), all candidates which are tied with some other candidate in radius r, are tied with c_1, and hence "can be tied with c_1 in $S_{d,r}(s)$" is a transitive relation. Since any candidate-wise metric is in particular neutral, $H_{d,r}$ is upward closed by the first part.

Let x be the lowest-ranked candidate participating in any tie. By upward-closeness, $(y,c_1) \in H_{d,r}$ for all y ranked weakly above x. Then by transitivity, any edge (y,z) where y,z, are ranked weakly above x is also in $H_{d,r}$, which means that $H_{d,r}$ is a clique. □

Proof of 2 is similar to Meir (2015), Lemma 2.

Theorem 13. *Suppose agents each have a cliqued epistemic model (not necessarily the same one). Then, iterative voting under Plurality must converge to OD equilibrium, from any initial state.*

Proof. For contradiction, let us assume the theorem is wrong and there is a cycle. That is, there is a sequence of scoring vectors (states) $s^1,\ldots s^q$ such that s^{j+1} is the outcome of an agent i making an OD move in s^j, and s^1 is the result of an agent making an OD move in s^q. Let B be the set of candidates whose score changes throughout the cycle, and let $z \in B$ be the candidate with the lowest score in the cycle (if there are multiple such candidates, let z be the lowest ranked in the tie-breaking rule).

Let $s^{q'}$ was be a state where z is at their lowest score, and in which an agent j makes a move, changing their vote from some candidate a to z. This means z was undominated at this point for j, which means all ties with B elements were within the same information set, and moreover, $z \succ^j c$ and $c \succ^j a$ for any $c \in B$ (Since the pivot-graph is a clique, there is a tie between each 2 candidates in B). However, as this is a cycle, there is a step $s^{\bar{q}}$, in which agent j changes their vote from $b \in B$ to a. This means b is dominated, and a is not, but this means there is some tie between a and another candidate x. Since $c \succ^j a$ for any $c \in B$, this means $x \notin B$.

If in $s^{q'}$ x's score was larger than z's, this means there was a tie between x and a was in the pivot-graph for agent j, and by moving to z, this indicates $x \succ^j a$. If x's score was smaller than z's in $s^{q'}$, the score of b in $s^{\bar{q}}$ is larger than that of x (since all scores are larger than that of z in $s^{q'}$), and since $b \succ_j a$, agent j should have preferred to stay with agent b. □

Theorem 14. *Suppose agents each have a concentric, cliqued epistemic model (not necessarily the same one). Then, iterative voting under Veto must converge to OD equilibrium, from any initial state.*

Proof. Assume, for contradiction, that the process does not converge. Let R be the set of candidates whose score changes an infinite number of times, and let $z \in R$ be the candidate which has the lowest score in the cycle (breaking ties using the tie-breaking rule), and let $\vec{s}^q$ be the state where it reaches this abysmal score. That is, some voter j moves from vetoing candidate a to vetoing candidate z. Candidate a's (and any other $c \in R$) score is above z's, as otherwise its own vetoing before would give it a lower score than z. Since this is a cliqued epistemic model, leaving a means it is the favorite candidate of voter j over all candidates with scores above z, in particular, for any $c \in R$, $a \succ^j c$.

At some point in the future $\vec{s}^{q'}$, due to the cycle, voter j will move from vetoing some candidate $b \in C$ to veto a, due to an edge in its relevant pivot-graph, indicating a tie between a and some other candidate x. If x's score at $\vec{s}^q$ was higher than z, then we know a is preferred over it from z's vetoing. If x's score was lower, we know it hasn't changed (as it isn't in R), meaning b is still tied with a as well in the pivot-graph of $\vec{s}^{q'}$ as it was in $\vec{s}^q$, hence voter j will not move (since $a \succ^j b$). □

Other convergence results from (Meir, Lev, and Rosenschein 2014) could be similarly extended for any upward-closed information structure. Theorem 14 with Prop. 12 imply the first non-Plurality result for local dominance.

Corollary 15. *Using any candidate-wise metric, local-dominance converges to an equilibrium when using Veto.*

Discussion and Future Directions

This paper presents a framework to model voting situations in which voters do not have perfect information of the world. Moreover, they do not even have an exact understanding of their uncertainty of the world's state. Hence, their understanding is modeled in a coarser way – as "shades of likelihood" of various voting outcomes, derived from a prospective poll. This framework is robust enough so as to allow us to capture many previously suggested heuristics and strategies of voter behavior under uncertainty. That is, we are able to express these heuristics as rational strategies under particular information structure known to players.

Indeed, the use of the pivot-graph and its topological properties to show convergence (or lack of it), opens the question of whether we can discuss issues of convergence in terms of graph structures (and the metrics or properties that induce them). The fact that ordinal dominance in a large population voting scenario can be computed efficiently, stands in contrast to the negative results in Conitzer et al. (2011), where verifying whether vote a' dominates a is NP-hard under the Borda rule. This is due to our simplifying assumption on the sharp pivot property that allows us to replace (arbitrarily complicated) information sets with a simple pivot graph representation.

A natural and important use of our model is to reformulate heuristics from various game-theoretic domains – not limited to social choice – as ordinally-dominant strategies. This might offer an insight into the built-in assumptions inherent in these heuristics, and allow, perhaps, novel formulations of new heuristics and methods, tailored to particular uncertainty structures.

Another promising direction is exploring possible connections between ordinal information structures and existing theories of qualitative uncertainty such as (Halpern 1997).

Acknowledgments

This research was partly supported by the Israel Science Foundation (ISF) under Grant No. 773/16 and by the Grant SUG M4082008.020. We acknowledge support from the UK EPSRC under Project EP/P031811/1 (Voting Over Ledger Technologies).

References

Boutilier, C. 1994. Toward a logic for qualitative decision theory. *KR* 94:75–86.

Brams, S. J., and Fishburn, P. C. 1978. Approval voting. *American Political Science Review* 72(3):831–847.

Chater, N.; Tenenbaum, J. B.; and Yuille, A. 2006. Probabilistic models of cognition: Conceptual foundations. *Trends in Cognitive Sciences* 10(7):287–291.

Conitzer, V.; Walsh, T.; and Xia, L. 2011. Dominating manipulations in voting with partial information. In *Proceedings of the 25th National Conference on Artificial Intelligence (AAAI)*, 638–643.

Desmedt, Y., and Elkind, E. 2010. Equilibria of plurality voting with abstentions. In *Proceedings of the 11th ACM conference on Electronic Commerce (EC)*, 347–356.

Dow, J., and Werlang, S. R. d. C. 1994. Nash equilibrium under knightian uncertainty: breaking down backward induction. *Journal of Economic Theory* 64(2):305–324.

Dubois, D.; Fargier, H.; and Prade, H. 1996. Refinements of the maximin approach to decision-making in a fuzzy environment. *Fuzzy Sets and Systems* 81(1):103–122.

Dutta, B., and Laslier, J.-F. 2010. Costless honesty in voting. in 10th International Meeting of the Society for Social Choice and Welfare, Moscow.

Elkind, E.; Markakis, E.; Obraztsova, S.; and Skowron, P. 2015. Equilibria of plurality voting: Lazy and truth-biased voters. In *International Symposium on Algorithmic Game Theory*, 110–122. Springer.

Grandi, U.; Loreggia, A.; Rossi, F.; Venable, K. B.; and Walsh, T. 2013. Restricted manipulation in iterative voting: Condorcet efficiency and Borda score. In *Proceedings of 3rd International Conference of Algorithmic Decision Theory (ADT)*, 181–192.

Halpern, J. Y. 1997. Defining relative likelihood in partially-ordered preferential structures. *Journal of Articial Intelligence Research* 7:1–24.

Laslier, J.-F. 2009. The leader rule: A model of strategic approval voting in a large electorate. *Journal of Theoretical Politics* 21(1):113–136.

Laslier, J.-F. 2010. Laboratory experiments on approval voting. In *Handbook on approval voting*. Springer. 339–356.

Matt, P.-A.; Toni, F.; and Vaccari, J. R. 2009. Dominant decisions by argumentation agents. In *International Workshop on Argumentation in Multi-Agent Systems*, 42–59. Springer.

Meir, R.; Lev, O.; and Rosenschein, J. S. 2014. A local-dominance theory of voting equilibria. In *Proceedings of the 15th ACM conference on Economics and Computation (EC)*, 313–330.

Meir, R. 2015. Plurality voting under uncertainty. In *Proceedings of the 29th Conference on Artificial Intelligence (AAAI)*, 2103–2109.

Meir, R. 2017. Iterative voting. In Endriss, U., ed., *Trends in Computational Social Choice*. AI Access. chapter 4, 69–86.

Myerson, R. B., and Weber, R. J. 1993. A theory of voting equilibria. *The American Political Science Review* 87(1):102–114.

Reijngoud, A., and Endriss, U. 2012. Voter response to iterated poll information. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, volume 2, 635–644.

Riker, W. H., and Ordeshook, P. C. 1968. A theory of the calculus of voting. *American political science review* 62(1):25–42.

Shafer, G. 1976. *A mathematical theory of evidence*, volume 42. Princeton university press.

Tversky, A., and Kahneman, D. 1974. Judgment under uncertainty: Heuristics and biases. *Science* 185(4157):1124–1131.

Williamson, T. 1992. Inexact knowledge. *Mind* 101(402):217–242.