Data Collection Prioritization: Choosing Observations to Contract Credal Sets

Michael Leon Khalfin

October 17, 2026

Abstract

Decisions often rest on quantities that the evidence does not determine. When the evidence is a set of constraints, the tightest bounds it supports on a quantity are the optima of two mathematical programs, and their difference, the width, measures what the evidence and the maintained assumptions leave open. We study data collection prioritization (DCP): choosing which costly questions to ask, within a budget, to minimize the worst-case widths of the quantities a decision depends on. We also treat scenario-structured credal sets with imprecise weights. At a fixed width, DCP is a weighted set covering problem over pairs of worlds that the design must separate. We solve it exactly by row generation with minimum-cardinality rows, and show that, unless P = NP, no polynomial-time method guarantees a width within any multiplicative factor of the optimum at the same budget. We formulate eleven applications, solve a food-web instance from real data to optimality, where buying one best question at a time makes no progress, and solve synthetic food webs to optimality up to 30 species.

1 Introduction

A school reports an outbreak. The school knows who shared a classroom with whom on each day, because it keeps timetables. It does not know who was infected, or when. A compartmental model in the tradition of Kermack and McKendrick [22] adds a few assumptions that nobody would call unreasonable: a person is infected at most once, is infectious for a bounded window that begins after a bounded delay, and can only be infected by someone who was infectious and in the room. Together with the timetable, these assumptions rule out most assignments of infection times to students. They do not rule out all of them. Many epidemics remain consistent with everything the school knows, and they disagree about the things the school needs to decide: how many students were infected after the first week, and which students could have infected many others.

Each such question has a smallest and a largest answer over the epidemics that the evidence still allows. If both answers are 12, the evidence and the school's assumptions determine the answer. If they are 3 and 41, they do not. These bounds are sharp under the assumptions the school maintains: both values are consistent with the evidence, and no value outside them is. The gap can be narrowed only by stronger assumptions, which the school would have to justify, or by data the school does not yet have. The school can buy some of that data. It can test a few students for antibodies, or ask a few families when symptoms began. It cannot test everyone. So it faces a concrete allocation problem: *which observations should it make, within its budget, so that the gap on the questions it cares about becomes as small as possible, whatever those observations turn out to say?*

This paper studies that problem in general. We call it *data collection prioritization* (DCP). The epidemic is our running example [23], and the rest of this section explains why the same problem

appears across operations research, statistics, and engineering, what it generalizes, and how to solve it.

1.1 Mathematical programs that infer

Operations research usually reads the feasible set of a mathematical program as the set of *decisions* available to an agent, and it optimizes a cost over that set. In this paper we read the feasible set differently, as the set of *worlds* that remain consistent with what is known: the facts, observations, bounds, and modeling assumptions in hand. In the school example a world is an assignment of infection times to students, and the admissible worlds are the assignments that satisfy the timetable and the transmission rules. The constraints are often linear with integer variables, because the evidence is logical (a contact did or did not happen) and the quantities are counts.

Given a quantity of interest, such as the number of late infections, we minimize it and maximize it over the same feasible set. Neither optimum is a recommendation. Together they bound what the evidence determines about the quantity. The imprecise-probability literature gives these two numbers a precise meaning [1, 37, 41]. If we refuse to assume anything about how likely the admissible worlds are relative to one another, the honest description of our beliefs is the set of all probability distributions on the admissible worlds, called the vacuous credal set. The minimum and maximum are the lowest and highest expectations of the quantity over that set, its lower and upper previsions. Their difference, the *width*, is zero exactly when the evidence, together with the assumptions, determines the quantity. Width measures how much the admissible worlds disagree, not the loss of a decision, so we always report it together with its two endpoints. When more is known, for instance that the worlds fall into scenarios whose weights are only partly known, the credal set is smaller and the previsions are tighter, but the recipe of solving two programs does not change.

Data enter as answers to questions. A question has a cost and an answer that depends on which world is true. Asking a set of questions and hearing the answers removes every world that would have answered differently, so the credal set contracts and, under the vacuous credal set, the width can only shrink. How much it shrinks depends on the answers, which are unknown when the questions are bought. DCP therefore chooses the questions within the budget that make the width smallest under the worst answers the evidence allows. Section 2 states the problem in full, including several quantities at once and adaptive policies that choose the next question after hearing the last answer.

1.2 How credal sets contract

Three established traditions already choose data to reduce uncertainty, and each, read through credal sets, contracts a set of distributions in its own way. *Active learning* asks for labels that split the version space, the set of hypotheses that agree with every label seen so far [8, 17, 31, 35]; the vacuous credal set on a version space is exactly the belief state of such a learner, but the usual goal is the whole hypothesis rather than bounds on a chosen quantity. *Bayesian experimental design* starts from a single prior and chooses the experiment that most reduces the expected spread of the posterior [5, 25]. Its credal set contains one distribution, so it has no width; it asks how to do well on average under one belief, not how far the evidence still leaves a quantity open. *Classical optimal design* chooses experiments that make a confidence region for a parameter small in a chosen sense [14, 24, 33]; a confidence region is again an admissible set. Classical design problems can be cast as DCP instances, by bounding the measurement errors jointly instead of giving them a distribution. DCP is broader, because its admissible set can be combinatorial and a single answer can rule out a

whole group of worlds. We leave a formal treatment of this connection to future work.

DCP combines three features that these traditions do not combine. It contracts the credal set only in the directions that move the quantities of interest, so the admissible set may stay enormous as long as those quantities are nearly determined. It plans the whole budget against the worst answers, which matters because questions can be worth little alone and a great deal together. And its admissible set can be the feasible set of any mathematical program, including a mixed-integer one, so combinatorial evidence such as timetables, network links, and logical rules enters directly instead of through a likelihood. The question that organizes this paper is therefore how credal sets contract under data, and how to choose data so that they contract where it matters.

1.3 From pairs of worlds to covering

Two worlds that give the same answer to every question asked cannot be told apart by the data, and the worst residual width is the largest difference in the quantity of interest between two such worlds (Lemma 4). Evaluating a design is therefore one optimization problem over two copies of the world, which is a mixed-integer program when the feasible set is mixed-integer linear. Choosing a design becomes a covering problem: to bring the width below a target, the design must, for every pair of worlds whose values differ by more than the target, ask at least one question that tells them apart. Infinitely many such pairs may exist, but a pair matters only through the set of questions that separate it, and there are finitely many such sets. Generating these covering constraints only when they are needed gives an exact algorithm for mixed-integer linear feasible sets (Section 3).

1.4 Contributions

Three bodies of work are closest to DCP. The test cover problem [30] is DCP with a finite admissible set and target width zero. Information-based complexity and optimal recovery [28, 36] define the diameter of information, which is the residual width of the pair identity, and show that adaptivity does not help for linear problems on convex, centrally symmetric sets. Robust optimization with decision-dependent information discovery [40] can express static DCP, with an interval as the decision taken after the data; its exact methods [32] cover a different class. Relative to these, the paper makes three contributions.

(i) It defines DCP for mathematical programs whose feasible sets describe admissible worlds, with lower and upper previsions as the quantities to bound and widths as the measure of what data leave open, first under the vacuous credal set (Section 2) and then under scenarios whose weights are only partly known (Section 6), where learning can widen a width and, for mixed-integer linear feasible sets, the inner problem is still one mixed-integer program.

(ii) It formulates eleven problems from distinct fields, among them traffic, power grids, ecology, genetics, and finance, as DCP instances (Section 4 and the appendix).

(iii) It gives an exact covering algorithm for mixed-integer linear feasible sets (Section 3) and runs every step of the method on a food web built from real gut-content records (Section 5), comparing the width a design guarantees with the width the revealed data actually leave.

Section 7 closes with limitations.

2 The problem class

This section states DCP under the weakest belief model, the vacuous credal set, in which the only thing we commit to is which worlds are possible. Section 6 adds structure on top of it. Throughout,

we follow one rule of interpretation: a constraint states what the evidence rules out, and an optimum gives the most extreme value of a quantity that the evidence still allows.

2.1 Information states and admissible worlds

Let $\Omega \subseteq \mathbb{Z}^{n_Z} \times \mathbb{R}^{n_R}$ be the ambient space of *worlds*. A world $\omega = (\omega_Z, \omega_R)$ is one complete specification of everything the model represents: in the epidemic, one infection time for every person; in a food web, one flow on every admissible link. An *information state* $\mathcal{I}$ collects the facts, observations, bounds, and modeling assumptions in hand. Its *admissible set* $\Omega_{\mathcal{I}} \subseteq \Omega$ is the set of worlds that $\mathcal{I}$ does not exclude. In our applications $\Omega_{\mathcal{I}}$ is mixed-integer linear,

$$\Omega_{\mathcal{I}} = \{\omega \in \Omega : A\omega \leq b,\ \omega_J \in \mathbb{Z}^{|J|}\}, \tag{1}$$

possibly after adding auxiliary variables that the constraints need. The theory below uses only the set $\Omega_{\mathcal{I}}$, so it applies to any admissible set. We assume $\Omega_{\mathcal{I}} \neq \emptyset$, since an empty admissible set means the evidence contradicts the model.

Information $\mathcal{I}'$ *refines* $\mathcal{I}$ when it rules out at least as much:

$$\mathcal{I}' \succeq \mathcal{I} \iff \Omega_{\mathcal{I}'} \subseteq \Omega_{\mathcal{I}}. \tag{2}$$

2.2 Credal sets and previsions

A quantity of interest is a function $f : \Omega \to \mathbb{R}$; following the imprecise-probability literature we also call it a *gamble* [41]. To speak about beliefs over worlds, we use a *credal set*: a set of probability distributions, each of which is compatible with what we know.

Definition 1 (Vacuous credal set)**.** The vacuous credal set of $\mathcal{I}$ is $\mathcal{K}^{\text{vac}}_{\mathcal{I}} = \{P \in \Delta(\Omega) : P(\Omega_{\mathcal{I}}) = 1\}$, the set of all probability distributions that put no mass outside the admissible set.

To avoid questions of integrability we take $\Delta(\Omega)$ to be the finitely supported distributions on Ω; none of the suprema and infima below changes.

The vacuous credal set says that a world outside $\Omega_{\mathcal{I}}$ is impossible and says nothing else. Inside $\Omega_{\mathcal{I}}$ a distribution may spread its mass however it likes. The lower and upper previsions of f are the smallest and largest expectations of f over the credal set,

$$\underline{E}_{\mathcal{I}}[f] = \inf_{P \in \mathcal{K}_{\mathcal{I}}} \mathbb{E}_P[f], \qquad \overline{E}_{\mathcal{I}}[f] = \sup_{P \in \mathcal{K}_{\mathcal{I}}} \mathbb{E}_P[f]. \tag{3}$$

Proposition 2 (Vacuous previsions are optima)**.** *Under the vacuous credal set,* $\underline{E}^{\text{vac}}_{\mathcal{I}}[f] = \inf_{\omega \in \Omega_{\mathcal{I}}} f(\omega)$ *and* $\overline{E}^{\text{vac}}_{\mathcal{I}}[f] = \sup_{\omega \in \Omega_{\mathcal{I}}} f(\omega)$.

Proof. Every $P \in \mathcal{K}^{\text{vac}}_{\mathcal{I}}$ satisfies $\mathbb{E}_P[f] \leq \sup_{\Omega_{\mathcal{I}}} f$, because P lives on $\Omega_{\mathcal{I}}$. Conversely, the point mass δ_ω belongs to $\mathcal{K}^{\text{vac}}_{\mathcal{I}}$ for every $\omega \in \Omega_{\mathcal{I}}$ and has expectation $f(\omega)$. The lower prevision follows in the same way. □

Proposition 2 is what allows a mathematical program to be used for inference. To bound what the evidence says about f, minimize and maximize f over $\Omega_{\mathcal{I}}$. The two optima are the *natural extension* of the evidence to f in the sense of Walley [41]: the tightest bounds that follow from the evidence without adding any assumption. Their difference

$$W_{\mathcal{I}}(f) = \overline{E}_{\mathcal{I}}[f] - \underline{E}_{\mathcal{I}}[f] \in [0, \infty] \tag{4}$$

is the *width* of f under $\mathcal{I}$. A width of zero means the evidence determines f. For the computational results we assume, as is true in every application of Section 4, that both optima are finite and attained. The theory allows $W = \infty$.

Events. An event $X \subseteq \Omega$ is a proposition about the world, such as "student i was infected after day F" or "the link from prey i to predator j exists." Its lower and upper probabilities are the previsions of its indicator, $\underline{P}_{\mathcal{I}}(X) = \underline{E}_{\mathcal{I}}[\mathbf{1}_X]$ and $\overline{P}_{\mathcal{I}}(X) = \overline{E}_{\mathcal{I}}[\mathbf{1}_X]$. Under the vacuous credal set,

$$\underline{P}^{\text{vac}}_{\mathcal{I}}(X) = \mathbf{1}\{\Omega_{\mathcal{I}} \subseteq X\}, \qquad \overline{P}^{\text{vac}}_{\mathcal{I}}(X) = \mathbf{1}\{\Omega_{\mathcal{I}} \cap X \neq \emptyset\}. \tag{5}$$

So the vacuous model sorts every proposition into three bins: *certain* (true in every admissible world), *impossible* (true in none), and *undecided* (true in some and false in others). The width $\overline{P} - \underline{P}$ is 1 exactly for the undecided propositions.

2.3 Metrics and objectives

A decision usually depends on more than one quantity. Let $u_1(\mathcal{I}), \ldots, u_m(\mathcal{I})$ be the widths that matter, each the width of a quantity of interest f_i or of an event X_i, and collect them in $u(\mathcal{I}) \in \mathbb{R}^m_+$. We call each u_i a *metric*. Given nonnegative weights w and a monotone norm $\|\cdot\|$, one that does not decrease when any component of a nonnegative vector increases, the *objective*

$$\Phi(\mathcal{I}) = \|w \odot u(\mathcal{I})\| \tag{6}$$

aggregates them into one number. The main examples are the ℓ_p norms, $1 \leq p \leq \infty$, which we write $\|\cdot\|_p$: with $p = 1$ the objective totals the weighted ambiguity, and with $p = \infty$ it reports the worst metric. The distinction between metrics and objectives is that a metric is something a stakeholder reads, whereas an objective is something the designer minimizes.

Why width, and its limits. Width measures how much the admissible worlds disagree about a quantity. It does not measure the loss from a downstream action. It is the right target when the decision turns on the value of the quantity. If an action is taken when f exceeds a threshold τ, the evidence leaves the action open exactly for the thresholds strictly between the lower and the upper prevision, a set whose length is the width, so a smaller width leaves fewer thresholds at which the action is undetermined. Width also has limits. It depends on the scale of each metric and on the weights and norm used to aggregate several metrics. It ignores where the interval lies: a width of zero means that all admissible worlds agree, which may be agreement that an outcome is bad. We therefore report the endpoints together with every width. An objective built on decision loss, such as the worst-case regret of the best action, is a different problem, and robust optimization with decision-dependent information discovery [40] handles such downstream recourse.

Decisions often concern a list of elements rather than a single number, for example which students to isolate. The natural quantity is then the number of elements whose status the evidence leaves open. Counting, in one world, the elements with a property and taking the width of that count does not measure this. Suppose the evidence says that exactly one of two students, A and B, was infected late. Every admissible world has one late infection, so the count has width zero, yet the school does not know whom to isolate. We therefore give each element its own copy of the world, a *witness world*, and ask of each element separately whether the evidence allows the property. In the example both students are undecided, and the witness-world count has width 2. In general the witness-world count is more pessimistic: its width is an upper bound on the width of the single-world count, because a sum of separate maxima is at least the maximum of the sum, and a sum of separate minima is at most the minimum of the sum. The next proposition shows that witness-world counts are a special case of (6) and need no separate theory.

Proposition 3 (Witness-world counts are sums of event widths)**.** *Let $X_1, \ldots, X_m$ be events and let $\bar{v} = \max \sum_e \mathbf{1}\{\omega^e \in X_e\}$ and $\underline{v} = \min \sum_e \mathbf{1}\{\omega^e \in X_e\}$, where each optimization ranges over*

independent copies $\omega^1, \ldots, \omega^m \in \Omega_{\mathcal{I}}$. *Under the vacuous credal set,* $\bar{v} = \sum_e \overline{P}_{\mathcal{I}}(X_e)$, $\underline{v} = \sum_e \underline{P}_{\mathcal{I}}(X_e)$, *and* $\bar{v} - \underline{v} = \sum_e (\overline{P}_{\mathcal{I}}(X_e) - \underline{P}_{\mathcal{I}}(X_e))$, *the objective* (6) *with* $p = 1$ *and unit weights over event widths.*

Proof. The copies share no constraints, so each optimization separates into m independent problems, and the e-th is $\max_{\omega^e \in \Omega_{\mathcal{I}}} \mathbf{1}\{\omega^e \in X_e\} = \overline{P}^{\mathrm{vac}}_{\mathcal{I}}(X_e)$ by (5), and likewise for the minimum. □

The proposition explains why the width of a witness-world count is exactly the number of undecided elements.

2.4 Questions and contraction

Data arrive as answers to questions. Let $\mathcal{A} = \{1, \ldots, M\}$ be the *question menu*. Each question a has an answer map $h_a : \Omega \to \mathcal{Z}_a$, which states what the answer would be in each world, and a cost $c_a \geq 0$. The answer map is the only link between the data and the model, and it need not describe what we traditionally think of as a measurement. A reading with bounded error fits this definition once the reading itself is part of the world: if question a reports a quantity $q(\omega)$ to within δ_a, the world carries the reported value r_a with $|r_a - q(\omega)| \leq \delta_a$, and $h_a(\omega) = r_a$. Two worlds are then told apart by their reports, not by their true values. A *design* $y \in \{0,1\}^{\mathcal{A}}$ selects questions; $\mathcal{Y}(B)$ denotes the designs with $c^\top y \leq B$ that also meet any other design restrictions.

Asking the questions in y and hearing the answers z leaves the worlds that would have answered the same way,

$$\Omega_{\mathcal{I}}(y, z) = \{\omega \in \Omega_{\mathcal{I}} : h_a(\omega) = z_a \text{ for every } a \text{ with } y_a = 1\}, \tag{7}$$

and $\mathcal{I}(y, z)$ denotes the refined information state with $\Omega_{\mathcal{I}(y,z)} = \Omega_{\mathcal{I}}(y, z)$. The answers that can actually occur are $\mathcal{Z}_{\mathcal{I}}(y) = \{z \in \prod_{a:y_a=1} \mathcal{Z}_a : \Omega_{\mathcal{I}}(y, z) \neq \emptyset\}$. Under the vacuous credal set, the updated credal set is again vacuous on the smaller set: conditioning any P that lives on $\Omega_{\mathcal{I}}$ on the event $\Omega_{\mathcal{I}}(y, z)$ gives a distribution that lives on $\Omega_{\mathcal{I}}(y, z)$, and every distribution on $\Omega_{\mathcal{I}}(y, z)$ arises this way.

Contraction is monotone. If $\mathcal{I}' \succeq \mathcal{I}$, then $W_{\mathcal{I}'}(f) \leq W_{\mathcal{I}}(f)$ for every quantity f and $\Phi(\mathcal{I}') \leq \Phi(\mathcal{I})$, because a maximum over a subset is no larger, a minimum over a subset is no smaller, and the objective is a monotone norm of a componentwise smaller vector. So, under the vacuous credal set, asking questions never increases any metric. This property is special to the vacuous credal set. Under the scenario credal sets of Section 6, an answer can widen the previsions (Example 11), and how credal sets contract becomes a real question.

2.5 Data collection prioritization

When we choose a design, we do not know the answers. DCP guards against the worst answers that some admissible world could produce. The *residual width* of a design is

$$W(y) = \sup_{z \in \mathcal{Z}_{\mathcal{I}}(y)} \Phi(\mathcal{I}(y, z)), \tag{8}$$

and the *static* DCP problem is

$$\min_{y \in \mathcal{Y}(B)} W(y). \tag{9}$$

An *adaptive* policy π chooses each question after hearing the previous answers. Formally, π maps each history $\tau = ((a_1, z_1), \ldots, (a_k, z_k))$ to "stop" or to a next question that does not already appear in τ. Since no question is repeated, every history has at most $|\mathcal{A}|$ questions, and a world ω and a

policy π generate a unique terminal history $\tau(\pi, \omega)$, the history at which π stops, because the world fixes every answer. We require that the questions asked along every terminal history form a design in $\mathcal{Y}(B)$. Prefixes of a terminal history need not be designs in $\mathcal{Y}(B)$ when $\mathcal{Y}(B)$ contains restrictions that are not preserved under removing questions. The budget is preserved under removing questions, because costs are nonnegative, so it holds along every prefix. The value of the policy is

$$V(\pi) = \sup_{\omega \in \Omega_{\mathcal{I}}} \Phi\big(\mathcal{I}(\tau(\pi, \omega))\big), \tag{10}$$

and the adaptive DCP problem minimizes $V(\pi)$. A static design is a policy that ignores answers, so the adaptive optimum never exceeds the static one. The ratio between them is the *adaptivity gap*.

The inner problem in (8) has the form of a max–min–max problem, but it reduces to a single maximization.

Lemma 4 (Pair identity)**.** *For a single quantity f under the vacuous credal set,*

$$W(y) = \sup\big\{f(\omega^+) - f(\omega^-) \;:\; \omega^+, \omega^- \in \Omega_{\mathcal{I}},\ h_a(\omega^+) = h_a(\omega^-)\ \forall a \text{ with } y_a = 1\big\}. \tag{11}$$

The same holds for the objective (6) *with $p = \infty$, with $f(\omega^+) - f(\omega^-)$ replaced by $\max_i w_i(f_i(\omega^+) - f_i(\omega^-))$, where event metrics use $f_i = \mathbf{1}_{X_i}$. For $p = 1$, $W(y)$ is the supremum of $\sum_i w_i(f_i(\omega_i^+) - f_i(\omega_i^-))$ over $2m$ worlds $\omega_1^{\pm}, \ldots, \omega_m^{\pm} \in \Omega_{\mathcal{I}}$, two per metric, that all give the same answers to the asked questions.*

Proof. For each $z \in \mathcal{Z}_{\mathcal{I}}(y)$, Proposition 2 gives $W_{\mathcal{I}(y,z)}(f) = \sup\{f(\omega^+) - f(\omega^-) : \omega^{\pm} \in \Omega_{\mathcal{I}}(y, z)\}$. Taking the supremum over z ranges over all pairs that give the same answers to the asked questions, since such a pair determines z and every such z is feasible. For $p = \infty$, the supremum over z commutes with the maximum over i. For $p = 1$, fix z: each metric's two worlds can be chosen separately in $\Omega_{\mathcal{I}}(y, z)$, so the largest total is the sum of the widths, and $2m$ worlds lie in a common $\Omega_{\mathcal{I}}(y, z)$ exactly when they all agree on the asked questions. □

For $p = 1$ the worlds of different metrics must agree with one another, not only within each pair, because the answer is shared by all metrics. Requiring agreement only within pairs would let one metric use one answer and another metric a different one, and would overstate the width. The inner problem is then a mixed-integer linear program over $2m$ copies of the world. Other norms, such as $p = 2$, lead to the same maximization over $2m$ copies but need their own encoding of the objective and a matching solver. In every case the static problem (9) is a min–max problem whose inner level is a mixed-integer program over copies of the world; our models link the copies to the answers with the constraints

$$h_a(\omega^\alpha) - z_a \le M_a(1 - y_a), \qquad z_a - h_a(\omega^\alpha) \le M_a(1 - y_a), \qquad \forall a \in \mathcal{A},\ \forall \alpha, \tag{12}$$

where α ranges over the copies, which is exact for bounded answers, integer or real, with M_a at least the range of h_a over $\Omega_{\mathcal{I}}$.

2.6 Four steps

In practice DCP is a four-step process, and we organize the applications and the computational study around it.

Step 1. *Formulate.* State the problem, explain why it is a program (what couples the unknowns and where integrality comes from), define the metrics and the question menu, and name the evidence that defines $\Omega_{\mathcal{I}}$.

Step 2. *Solve the inner problem at scale.* Compute previsions and widths. An off-the-shelf solver may be enough; when it is not, the structure of $\Omega_{\mathcal{I}}$ suggests a custom method.

Step 3. *Design the policy.* Solve the outer problem (9) or its adaptive version.

Step 4. *Interpret.* Report what is certain, what is impossible, what is undecided, and what the chosen data would settle.

2.7 The running example

In the epidemic model [23], the world is a vector of infection times, one per person, each in the time horizon or "never." The admissible set is defined by the contact data and four rules: no reinfection, a fixed number of seeds at the start, an infectious window $[L, U]$ after infection, and a transmission rule requiring every non-seed infection to follow an infectious contact within a delay window. The model writes these rules with binary infection, infectiousness, and transmission variables, and all but the infection times are definitional (Appendix A.1). Two of the four metrics are single-world (the number infected after day F, the number of superspreaders after day F) and two are witness-world counts (who could be infected after day F, and who could be a superspreader), which by Proposition 3 are sums of event widths. A question asks whether person i was infected during time bin m, with a binary answer, and the budget limits how many such questions we ask.

2.8 Notation

Table 1 collects the core notation of DCP. Notation used only in a later section, or inside one application, is defined where it appears.

Table 1: Notation.

Symbol	Meaning	Where
ω, Ω	a world (one complete specification of the system); the space of worlds	§2.1
$\mathcal{I}$, $\Omega_{\mathcal{I}}$	information state; its admissible set, the worlds it does not rule out	§2.1
$\mathcal{I}' \succeq \mathcal{I}$	$\mathcal{I}'$ refines $\mathcal{I}$: $\Omega_{\mathcal{I}'} \subseteq \Omega_{\mathcal{I}}$	(2)
$\mathcal{K}_{\mathcal{I}}$, $\mathcal{K}_{\mathcal{I}}^{\text{vac}}$	credal set; the vacuous credal set on $\Omega_{\mathcal{I}}$	Def. 1
f, X	a quantity of interest (gamble); an event	§2.2
$\underline{E}_{\mathcal{I}}[f]$, $\overline{E}_{\mathcal{I}}[f]$	lower and upper previsions of f	(3)
$\underline{P}_{\mathcal{I}}(X)$, $\overline{P}_{\mathcal{I}}(X)$	lower and upper probabilities of X	(5)
$W_{\mathcal{I}}(f)$	width of f: $\overline{E}_{\mathcal{I}}[f] - \underline{E}_{\mathcal{I}}[f]$	(4)
$u(\mathcal{I})$, w, $\Phi(\mathcal{I})$	vector of metrics (widths); weights; objective $\|w \odot u(\mathcal{I})\|$	(6)
$\mathcal{A}$, a	question menu; a question	§2.4
h_a, $\mathcal{Z}_a$, c_a	answer map of a; its answer set; its cost	§2.4
y, $S(y)$, $\mathcal{Y}(B)$	design (which questions to ask); its set of questions; designs within budget B	§2.4
z, $\Omega_{\mathcal{I}}(y, z)$, $\mathcal{Z}_{\mathcal{I}}(y)$	answers; worlds that agree with them; answers that can occur	(7)
$W(y)$	residual width: worst case of Φ over answers	(8)
π, τ, $V(\pi)$	adaptive policy; history of questions and answers; value of a policy	(10)
ω^+, ω^-	the two worlds of the pair problem ($2m$ worlds $\omega_i^{\pm}$ for $p = 1$)	(11)
M_a	big-M constant in the linking constraints	(12)