Cascading Failures:
A Preliminary Info-Gap Analysis

Yakov Ben-Haim
Yitzhak Moda'i Chair in Technology and Economics
Technion — Israel Institute of Technology
Haifa, Israel
yakov@technion.ac.il

Contents

Abstract Cascading failures occur in networks of interacting agents in which failure at one node can cause further failures. We define the 'degree of cascading failure' as the fraction of nodes that *could* fail as a result of at least one single failure. This refers to the *possibility* of failure, and thus involves the *uncertainty* of failures. We employ the concept of robustness, as developed in info-gap theory, to study cascading failures. This method attempts to satisfy an outcome requirement, and to maximize the robustness against uncertainty. This is a procedural optimization rather than a substantive optimization. The outcome is the substantive "good" that one seeks, and we only attempt to make it good enough, while the robustness is an aspect of the procedure of reaching a decision, and the robustness is optimized.

We illustrate three properties of the robustness analysis. 'Zeroing' asserts that predicted outcomes (e.g. degree of cascading failure) have no robustness against uncertainty. 'Trade off' asserts that greater robustness is obtained only in exchange for more modest outcome goals (e.g. accepting greater degree of cascading failure). 'Preference reversal' between policy alternatives arises in situations where one alternative is putatively better but more uncertain. This presents the policy maker with a dilemma: choose the putatively better but more uncertain option, or choose the putatively worse but more reliable one? The info-gap robustness analysis offers a resolution of this dilemma. This analysis underlies a critique of conventional optimization in which one uses the

\papers\emergency-readiness\workshop2018\cas-fail003.tex. 14.11.2018 best data, knowledge and understanding to prioritize the decision alternatives based on predicted outcomes.

We use two simple examples of static networks to explore the implications of deep uncertainty, and to demonstrate the analysis of robustness based on info-gap theory. The first example is a general linear network with uncertain probability of successful transmission from one node to the next. The second example is the analysis of traffic jams as a cascading failure problem. In both examples we explore how the analysis supports design and policy decisions.

1 Introduction

Fraid (2016) has studied the application of info-gap decision theory to cascading failures in dynamic networks involving supply and demand requirements at each node. In this paper we will use simple static examples to illustrate the insights obtained from an info-gap robustness analysis of cascading failures in networks.

Info-gap theory is a method for prioritizing options and making choices and decisions under severe uncertainty (Ben-Haim, 2006, 2010, 2018). The options might be operational alternatives (implement a policy, choose a budget, decide to intervene or not, etc.) or more abstract decisions (choose a model structure, make a forecast, formulate a policy, etc.). Decisions are based on data, scientific theories, empirical relations, knowledge and contextual understanding, all of which we'll refer to as one's *models,* and these models often recognize and quantify uncertainty.

Info-gap theory has been applied to decision problems in many fields, including various areas of engineering (Kanno and Takewaki, 2006; Chinnappen-Rimer and Hancke, 2011; Harp and Vesselinov, 2013), biological conservation (Burgman, 2005), economics (Knoke, 2008; Ben-Haim, 2010), medicine (Ben-Haim *et al.,* 2012), national security (Moffitt, Stranlund, and Field, 2005), public policy (Hall *et al.,* 2012), and more (info-gap.com). Info-gap robust-satisficing has been discussed non-technically elsewhere (Schwartz, Ben-Haim, and Dacso, 2011; Ben-Haim, 2012 a, b; Ben-Haim, 2018).

In section 2 we define and illustrate the concept of the degree of cascading failure of a network. In section 3 we explain the centrality of uncertainty in evaluating the degree of cascading failure of a linear network with uncertain transmission probabilities between nodes. We then apply the info-gap analysis of robustness to manage this uncertainty. In section 4 we illustrate the methodology in managing the cascading failure of highway traffic jams. On a more general level, our analysis is a critique of conventional optimization, as explained in the concluding discussion, section 5.

2 Degree of Cascading Failure: A Definition

Consider a network with n nodes. The network has **cascading failures of degree** ϕ if there is at least one node in the network whose failure *could* cause the failure of at least a fraction ϕ of the nodes in the network.

Example 1 For example, consider the game of 'telephone' in which n people sit in a row, and one person whispers a message to their neighbor, who whispers the message to the next neighbor, etc. The last person to receive the message announces it out loud. That announcement often differs greatly from the original version, to everyone's delight. This game involves n transmissions of the message (including the final announcement), any one of which could corrupt the message. We will assume that no corruption could correct a previous corruption.[1] A node fails if the message that it transmits is a corruption of the original message. The first transmission could be corrupted, meaning that all n transmissions of nodes $1, \ldots, n$ would be corrupted. Thus this is a network with cascading failure of degree $\phi = 1$. ∎

Example 2 Consider a central broadcaster, node 1, who transmits a message that is received by $n-1$ receivers, nodes $2, \ldots, n$. The broadcaster fails if its transmission is faulty. Any of the nodes $2, \ldots, n$ fails if it corrupts the message after receipt. A receiver's corruption may or may not correct for corruption by the original transmission. If node 1 fails, then all other nodes could also fail, meaning that this network has cascading failure of degree $\phi = 1$. ∎

Example 3 Consider n observers watching the same scene, for instance a volcano. Assume that this scene is in one of two well defined states. For instance, either "eruption" or "no eruption" of the volcano. Observer i fails if it records the scene incorrectly. There is no communication between observers, so failure of any one observer has no effect on any other observer. Hence this network has cascading failure of degree $\phi = 1/n$. ∎

Example 4 Continue the previous example, but assume that n is an even number of nodes. Also, each node discusses its observation with exactly 1 other node, and they both then agree on the observation. In other words, any node could corrupt both itself and at most 1 other node. Thus this network has cascading failure of degree $\phi = 2/n$. ∎

Example 5 Continue the previous example except that now node 1 discusses with node 2 and they agree on their observation. Then node 2 discusses with node 3 and they agree, and so on. Thus, failure by node 1 could result in failure of all nodes. Thus this is a network with cascading failure of degree $\phi = 1$. ∎

Example 6 *Hierarchical network.* Consider a triangular hierarchical system such as shown in fig 1. Each node on the bottom row has received a message that it passes to either 1 or 2 nodes in the layer just above. Every receiving node does the same. Each node may corrupt the message it has received. A single node in the bottom row of the 2-row network could pass a corrupted message to the top node so the degree of cascading failure of this 3-node network is $\phi = 2/3$. In the 3-row network with 6 nodes, the central node in the bottom row could "infect" three higher nodes, so the degree of cascading failure is $\phi = 4/6 = 2/3$. As shown in appendix A, the degree of cascading failure in a triangular hierarchical network with either $2n$ or $2n+1$ rows is:

$$\phi_{2n} = \phi_{2n+1} = \frac{n+1}{2n+1}, \quad n = (0,)\, 1,\, 2,\, \ldots \tag{1}$$

where the expression for ϕ_{2n+1} holds also when $n = 0$.

It is interesting to note that the degree of cascading failure decreases as the size of the hierarchy increases:

$$\phi_{2n} > \phi_{2(n+1)} \tag{2}$$

[1]For example, if the message is a single binary bit, a corruption would change the bit, and a subsequent corruption would change the bit back to its original value, which thus corrects the previous corruption. We exclude this possibility by considering only complex messages.

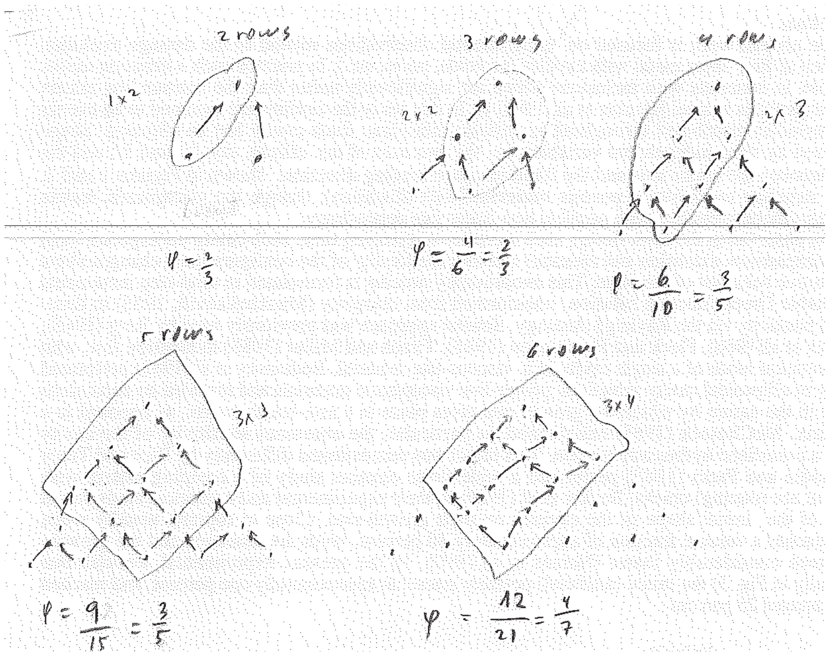

Figure 1: Five triangular hierarchical networks for example 6.

The degrees of cascading failure for networks with $1, 2, , \ldots, 6$ rows are 1, 2/3, 2/3, 3/5, 3/5, and 4/7, respectively.

Furthermore, the degree of cascading failure converges asymptotically:

$$\lim_{n\to\infty} \phi_{2n} = \frac{1}{2} \tag{3}$$

The convergence to a degree of 1/2 is fairly rapid. ∎

3 Linear Network with Uncertain Transmission Probability

3.1 Cascading Failures and Uncertainty

The concept of degree of cascading failure refers to the *possibility* that failure of a single node could cause additional nodes to fail. Thus, the degree of cascading failure involves the *uncertainty* of failures.

For instance, consider the game of 'telephone' discussed earlier in example 1. A node may fail either by receiving and accurately transmitting a corrupted message, or by corrupting and then transmitting the message. If the first node corrupts the message, then all subsequent nodes transmit corrupted messages either by accurately transmitting the corrupted message or by further corrupting the message. Because the first node *could* fail, the degree of cascading failure of the 'telephone' network is $\phi = 1$.

Let π denote the probability that any single node accurately transmits the message it received. We will assume that the probability of accurately transmitting the received message is the same for all nodes, and that correct transmission is statistically independent between nodes.

The degree of cascading failure, ϕ, is not a random variable; it is a deterministic property of the network: the largest possible fraction of nodes that could fail in a cascading failure. However, the fraction of nodes involved in a cascading failure can be smaller than ϕ. Let ψ denote the fractional size of a cascading failure in a network with n nodes. Thus ψ is a random variable with values $0/n,\ 1/n,\ 2/n,\ \ldots,\ n/n$. As shown in appendix B, the probability distribution of ψ, for the game of 'telephone', is:

$$P\left(\psi = \frac{i}{n}\right) = \begin{cases} \pi^n, & \text{if } i = 0 \\ \pi^{n-i}(1-\pi), & \text{if } i = 1,\ \ldots,\ n \end{cases} \tag{4}$$

The cumulative probability distribution of ψ is:

$$P\left(\psi \le \frac{i}{n}\right) = \pi^{n-i}, \quad i = 0,\ 1,\ \ldots,\ n \tag{5}$$

The probability of cascading failures larger than i/n is the complementary probability:

$$P\left(\psi > \frac{i}{n}\right) = 1 - \pi^{n-i}, \quad i = 0,\ 1,\ \ldots,\ n \tag{6}$$

At fixed network size, n, and fixed probability π, we see that the probability of cascading failures larger than i/n decreases as i/n increases. This is not saying that larger cascading failures are less likely than smaller cascading failures. Rather, the set of all cascading failures gets smaller as i/n increases. Thus the probability of cascading failures larger than i/n decreases as i/n increases.

3.2 Cascading Failures and Robustness

We suppose that we have an estimate, $\widetilde{\pi}$, of the probability of accurate transmission from one person to the next in the game of 'telephone', and an estimated error of this estimate, a positive number s. Roughly speaking, π is estimated as $\widetilde{\pi} \pm s$, where we recognize that the actual error may exceed s. More precisely, the fractional error of π with respect to $\widetilde{\pi}$ is $|\pi - \widetilde{\pi}|/s$, whose magnitude is unknown. Thus the info-gap model of uncertainty is:

$$\mathcal{U}(h) = \left\{\pi:\ \pi \in [0,1],\ \ \left|\frac{\pi - \widetilde{\pi}}{s}\right| \le h\right\}, \quad h \ge 0 \tag{7}$$

Like all info-gap models of uncertainty, this is an unbounded family of nested sets.[2] When $h = 0$ the set contracts to the estimated value: $\mathcal{U}(0) = \{\widetilde{\pi}\}$. As h increases the sets become more inclusive. Thus h is called the horizon of uncertainty, and its value is unknown.

We would like to design or manage the network so that the probability of large cascades is small. More precisely, we require that the probability of cascading failures with degree larger than i/n have probability no greater than a critical value $P_{\rm c}$. That is, we require:

$$P\left(\psi > \frac{i}{n}\right) \le P_{\rm c} \tag{8}$$

For any choice of this critical probability, $P_{\rm c}$, we would like to know if this requirement is feasible. Alternatively, we would like to know what $P_{\rm c}$ values are feasible. The answers to these questions is provided by the robustness function.

The robustness is the greatest horizon of uncertainty, h, in the info-gap model of eq.(7), up to which the requirement in eq.(8) is guaranteed. Formally, the robustness to uncertainty in π is defined as:

$$\widehat{h}(P_{\rm c}, i) = \max\left\{h:\ \left(\max_{\pi \in \mathcal{U}(h)} P\left(\psi > \frac{i}{n}\right)\right) \le P_{\rm c}\right\} \tag{9}$$

[2]The family of sets is unbounded in the space of possible probability values, $[0,\ 1]$.

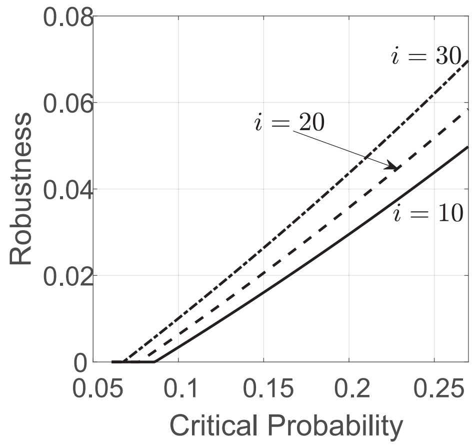

Figure 2: Robustness curves for 3 values of i. $\tilde{\pi} = 0.999$, $n = 100$, $s = 0.05$.

As shown in appendix C, the robustness function is:

$$\widehat{h}(P_{\rm c}, i) = \frac{1}{s}\left(\tilde{\pi} - (1 - P_{\rm c})^{1/(n-i)}\right) \tag{10}$$

or zero if this is negative.

The best-estimated prediction of the probability that the degree of cascading failure exceeds i/n is, from eq.(6), $P(\psi > i/n) = 1 - \tilde{\pi}^{n-i}$. From eq.(10) we see that the robustness reaches the horizontal axis precisely when the critical probability, $P_{\rm c}$, equals this value. That is, if $1 - \tilde{\pi}^{n-i}$ is adopted as the performance requirement, $P_{\rm c}$ in eq.(8) — and some would consider this sensible because it is the best estimate of the probability that the degree exceeds i/n — then the robustness for satisfying this requirement is zero. This is the **zeroing property:** best-model predicted outcomes have no robustness against uncertainty underlying the predictions.

Eq.(10) also shows the irrevocable **trade off** between robustness and performance: as the performance requirement becomes more demanding (as $P_{\rm c}$ is decreased) the robustness becomes smaller ($\widehat{h}(P_{\rm c}$, i) goes down). This is a trade **off** because we would like $P_{\rm c}$ to be small and $\widehat{h}$ to be large.

The robustness function of eq.(10) is shown in fig. 2 for $n = 100$, $\tilde{\pi} = 0.999$, $s = 0.05$, and three different values of i: 10, 20 and 30. As we expect from eq.(10) and the zeroing property, the robustness vanishes when $P_{\rm c} = 1 - \tilde{\pi}^{n-i}$, which equals 0.087, 0.078 and 0.068 for $i =$ 10, 20 and 30, respectively. Furthermore, the robustness increases as $P_{\rm c}$ increases as seen from the positive slopes of the curves, expressing the trade off property. We also see that the robustness increases as i increases. This is because the probability of cascading failures larger than i/n decreases as i/n increases, as explained following eq.(6).

We now consider two different implementations of the telephone game, both with $\tilde{\pi} = 0.999$. In one we have better knowledge so its error estimate, s, is smaller, but the number of players is greater, so n is larger. In both cases $i/n = 0.2$. Specifically, in the first implementation, $n_1 = 100$, $s_1 = 0.05$ and $i = 20$ as before (middle curve of fig. 2), and the estimated probability that the degree of cascading failure exceeds i/n is $P(\psi > i/n) = 0.078$. In the second implementation $n_1 = 60$, $s_1 = 0.15$ and $i = 12$ and $P(\psi > i/n) = 0.048$. That is, the poorer information in the second case ($s_2 > s_1$) is compensated by the smaller number of players ($n_2 < n_1$).

Plots of robustness curves for these two implementations are shown in fig. 3. The robustness

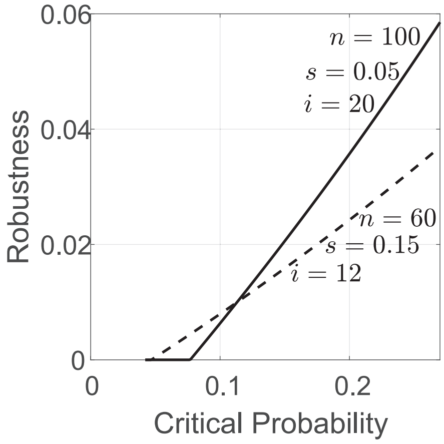

Figure 3: Robustness curves for 2 values of n, s and i. $\widetilde{\pi} = 0.999$.

curves cross one another, implying the potential for a **reversal of preference** between these two implementations. The second implementation (dashed curve) is putatively better: its predicted probability that ψ exceeds i/n is lower because of the smaller size of the network, as we expect from eq.(6). However, from the zeroing property we know that predicted outcomes have no robustness to uncertainty. From the trade off property we know that only poorer outcomes (larger P_c) have greater robustness to uncertainty. However, the trade off is more severe for the putatively better case (dashed curve) because the uncertainty, as expressed by s_2, is greater. As a consequence, the robustness curves cross one another at $P_c = 0.11$. If one needs a critical probability less than 0.11, then option 2 is more robust, though its robustness is small. If greater P_c would be acceptable then option 1 is more robust.

4 Traffic Jams and Network Management

High-speed freeways have been known to experience vast traffic jams in the middle of nowhere. These traffic jams arise from the relation between car speed and traffic density: speed decreases as density increases. The density of cars is low upstream of the jam, so car speeds are high. When some random disturbance causes a local increase of density, the local speeds drop. However, cars continue arriving at the disturbance at high speed, causing further increase in density and decrease in speed. The cascading failure here is that the length of the traffic jam grows and the density of cars increases, ultimately causing complete stoppage of car motion. The jam “evaporates” at the free leading edge at a low rate, while the jam grows at the upstream edge at the high rate of arrival of high-speed cars.

Such traffic jams could be ameliorated by automatic sensor detection of small jams, followed by automatic change of the maximum permitted speed upstream of the jam, fed back to upstream drivers by digital speed signs. The upstream speeds remain much greater than speeds at which cars leave the jam downstream, but the upstream density is much lower, which can allow the jam in contract and eventually vanish. We will study this as a network-management problem, where the basic question we address is determination of the revised upstream speed once a traffic jam is detected.

4.1 Traffic Model

Traffic behavior is a complex interaction of a large number of independent agents: the drivers. The drivers adjust their behavior (speed) in response to the density of cars in their vicinity, this adjustment further modifies the local density, and these modifications propagate along the freeway. We will adopt a simple macroscopic model of this multi-agent network. A more realistic analysis would require a dynamic multi-agent model.

Let x km denote the position along the freeway, let $\rho(x)$ cars/km denote the density of cars at position x, and let $s(x)$ km/hr denote the speed of cars at position x. The flux of cars at position x is the number of cars passing that position per hour, denoted $f(x)$ cars/hr, where $f(x) = \rho(x)s(x)$.

We consider a traffic jam in which cars are totally stopped along a length of freeway between positions $x = -x_\mathrm{j}$ and $x = 0$. We make the simplifying assumptions that car speeds are constant (and high) upstream of the jam, and zero in the jam. Likewise, density is constant (and low) upstream and constant (and high) in the jam. Specifically car speeds are:

$$s(x) = \begin{cases} s_0, & \text{if } x \leq -x_\mathrm{j} \\ 0, & \text{if } -x_\mathrm{j} < x \leq 0 \end{cases} \tag{11}$$

where s_0 is a positive number, like 100 km/hr. Likewise, the density of cars is:

$$\rho(x) = \begin{cases} \rho_0, & \text{if } x \leq -x_\mathrm{j} \\ \rho_1, & \text{if } -x_\mathrm{j} < x \leq 0 \end{cases} \tag{12}$$

where ρ_0 and ρ_1 are positive numbers, like 20 cars/km and 300 cars/km, respectively.

Consider the cars at the free downstream edge of the traffic jam. The density of cars immediately in front of these cars is zero (or very low) so they accelerate strongly, rapidly reaching a low speed, s_d, while detaching themselves from the jam. If the density of cars during this initial low-speed detachment stays essentially constant at ρ_1, then the flux of cars departing from the jam is $f_0 = \rho_1 s_\mathrm{d}$. For instance, if $s_\mathrm{d} = 10$ km/hr and $\rho_1 = 300$ cars/km, then $f_0 =$ 3000 cars/hr.

Let $N_\mathrm{j}(0)$ denote the number of stopped cars in a traffic jam at some initial reference time. The initial length of the traffic jam is x_j, so $N_\mathrm{j}(0) = \rho_1 x_\mathrm{j}$. The number of cars changes over time, evolving as $N_\mathrm{j}(t)$, because cars arrive from upstream and depart downstream. $N_\mathrm{j}(t)$ grows due to the flux $\rho_0 s_0$ of arriving cars, and shrinks due to the flux $\rho_1 s_\mathrm{d}$ of detaching cars. Thus:

$$\frac{\mathrm{d}N_\mathrm{j}(t)}{\mathrm{d}t} = \rho_0 s_0 - \rho_1 s_\mathrm{d} \tag{13}$$

The solution of this differential equation is:

$$N_\mathrm{j}(t) = \rho_1 x_\mathrm{j} + (\rho_0 s_0 - \rho_1 s_\mathrm{d})t \tag{14}$$

The time required for the jam to disappear, denoted T_d, is the value of t at which $N_\mathrm{j}(t) = 0$:

$$T_\mathrm{d} = \frac{\rho_1 x_\mathrm{j}}{\rho_1 s_\mathrm{d} - \rho_0 s_0} \tag{15}$$

provided that this is non-negative. If this expression is negative, then the jam is growing faster than it is shrinking, and it will never disappear. The value of T_d with the values specified earlier, for a jam 1 km long, is 0.3 hr.

4.2 Uncertainty and Robustness

We will study the modelling and management of uncertainty in the time required for the traffic jam to disappear, eq.(15). A jam of length $x_{\rm j}$ has been detected, the density of cars in the jam is ρ_1 and the speed of departing cars is $s_{\rm d}$. Before intervention, the density and speed of cars upstream of the jam are ρ_{0b} and s_{0b}. These values are all known.

The upstream maximum speed will be revised to a new lower value, $\widetilde{s}_{0r}$. However, the actual speed of upstream cars, denoted s_{0r}, may differ and may be greater. Our estimate is that $\widetilde{s}_{0r}$ may err by as much as $\pm w_s$ or more.

The upstream density will change as a result as the revised speed limit, the new value being ρ_{0r}. The relation between density and speed is poorly understood. Our best available model is linear:

$$\frac{{\rm d}\rho_{0r}}{{\rm d}s_{0r}} = -\delta \tag{16}$$

where δ is estimated to take the value $\widetilde{\delta}$, but this may err by $\pm w_\delta$ or more. From eq.(16) we can write:

$$\rho_{0r} = -(s_{0r} - s_{0b})\delta + \rho_{0b} \tag{17}$$

After revising the upstream speed limit, the duration of the jam, eq.(15), is:

$$T_{\rm d} = \frac{\rho_1 x_{\rm j}}{\rho_1 s_{\rm d} - \rho_{0r} s_{0r}} \tag{18}$$

where both ρ_{0r} and s_{0r} are uncertain.

Combining eqs.(17) and (18), the duration depends on the uncertain quantities, s_{0r} and δ, as:

$$T_{\rm d}(s_{0r}, \delta) = \frac{\rho_1 x_{\rm j}}{\rho_1 s_{\rm d} - [(s_{0b} - s_{0r})\delta + \rho_{0b}] s_{0r}} \tag{19}$$

The uncertainty in s_{0r} and δ is represented by the following fractional-error info-gap model:

$$\mathcal{U}(h) = \left\{ s_{0r}, \delta : \ s_{0r} \ge 0, \ \left| \frac{s_{0r} - \widetilde{s}_{0r}}{w_s} \right| \le h, \ \left| \frac{\delta - \widetilde{\delta}}{w_\delta} \right| \le h \right\}, \quad h \ge 0 \tag{20}$$

Note that the info-gap model depends on the revised upstream speed limit, $\widetilde{s}_{0r}$.

We require that the duration of the jam not exceed a critical value, $T_{\rm c}$:

$$T_{\rm d}(s_{0r}, \delta) \le T_{\rm c} \tag{21}$$

The robustness to uncertainty, or any choice $\widetilde{s}_{0r}$ of the revised upstream speed limit, is the greatest horizon of uncertainty, h, up to which all realizations of s_{0r} and δ in the uncertainty set $\mathcal{U}(h)$, satisfy the requirement in eq.(21). Formally:

$$\widehat{h}(T_{\rm c}, \widetilde{s}_{0r}) = \max \left\{ h : \ \left(\max_{s_{0r}, \delta \in \mathcal{U}(h)} T_{\rm d}(s_{0r}, \delta) \right) \le T_{\rm c} \right\} \tag{22}$$

Before evaluating robustness curves, let's consider some empirical data. In California, the Engineering and Traffic Survey typically sets speed limits near the 85th percentile speed,[3] which is nearly 1 standard deviation above the mean of a normal distribution. A typical highway may have an

[3] Division of Traffic Operations, California Department of Transportation, 2014, *California Manual for Setting Speed Limits,* p.14. http://www.dot.ca.gov/trafficops/camutcd/docs/california-manual-for-setting-speed-limits.pdf. Accessed online 9 November 2018.

average speed of about 80 km/hr, with a speed limit of 100 km/hr at the 85th percentile, implying a standard deviation of about 20 km/hr. A speed of 90 km/hr is at the 69th percentile, and 110 km/hr is at the 93rd percentile, so characterizing extremal speeds as '100 $\pm$10 km/hr or more' is a realistic description of the high-end drivers.

Fig. 4 shows a robustness curve based on eq.(22). The parameter values for this calculation are as follows. The traffic jam is of length $x_{\rm j} = 1$ km, the car density in the jam is $\rho_1 = 300$ cars/km, and the speed at which cars depart from the jam is $s_{\rm d} = 10$ km/hr. The initial upstream car density is $\rho_{0b} = 20$ cars/km, and the initial upstream speed is $s_{0b} = 100$ km/hr. The revised speed limit is $\widetilde{s}_{0r} = 90$ km/hr where the error estimate for observance of this limit is $w_{\rm s} = 10$ km/hr or more. The slope of density vs. speed is estimated as $\widetilde{\delta} = 0.1$ car hr/km^2, for which the predicted upstream car density is $\widetilde{\rho}_{0r} = 21$ cars/km. The estimated error of $\widetilde{\delta}$ is $w_\delta = 0.03$. The putative predicted duration of the traffic jam, after revising the speed limit, is $T_{\rm d} = 0.27$ hr.

In fig. 4 we observe the zeroing and trade off properties of all info-gap robustness curves. The robustness for achieving the predicted traffic jam duration of 0.27 hr is precisely zero, and robustness is increased only by accepting greater duration of the jam.

A substantial reduction in the speed limit should decrease the duration of the traffic jam. For instance, in a situation with mean and standard deviation of 80 and 20 km/hr, respectively, a speed of 70 km/hr is at the 31st percentile. One should expect consider violation of this revised speed limit, and, in the revised situation, characterizing extremal speeds as '70 $\pm$40 km/hr or more' is a reasonable description of the high-end drivers.

If the speed limit were reduced to $\widetilde{s}_{0r} = 70$ km/hr then the predicted upstream car density would be greater, $\widetilde{\rho}_{0r} = 23$ cars/km, but the predicted jam duration would still be less: $T_{\rm d} = 0.22$ hr. However, this more restrictive speed limit would be expected to induce greater speed violation, so the estimated error is $w_{\rm s} = 40$ or more. The dilemma is that the lower speed limit has a putatively lower jam duration but this is more uncertain. Fig. 5 shows the robustness curve for this option, together with the option shown in fig. 4. The robustness curves cross one another, indicating the potential for a reversal of preference between these options.

The robustness curves in fig. 5 cross one another at a traffic jam duration of 0.29 hr. If a duration less than this value is required, then the revised speed of $\widetilde{s}_{0r} = 70$ km/hr is more robust and hence preferred over $\widetilde{s}_{0r} = 90$, though the robustness is not large. If a greater jam duration is acceptable, then $\widetilde{s}_{0r} = 90$ is more robust and hence preferred. The choice between these options is a dilemma: $\widetilde{s}_{0r} = 70$ is putatively better but more uncertain than $\widetilde{s}_{0r} = 90$. The dilemma is manifested in the intersection of the robustness curves, which also supports a choice between these options.

5 Discussion

We have used the term "models" to refer to one's data, knowledge and understanding. The conventional optimization of a decision begins by using the best models that one has to predict the outcomes of the decision alternatives. One then chooses the option whose predicted outcome is best. This "best-model optimization" makes sense when one's models are fairly good. However, when the models are subject to deep uncertainty, then the zeroing property of the info-gap robustness function demonstrates that the predictions have no immunity against this uncertainty: their robustness is precisely zero. This means that the prioritization of the decision alternatives, based on their best-model predicted outcomes, is unreliable.

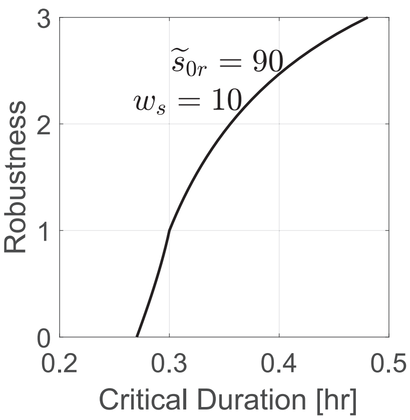

Figure 4: Robustness curve for revised upstream speed limit.

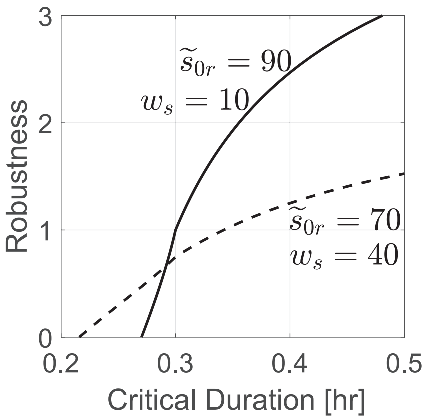

Figure 5: Robustness curves for two revised upstream speed limits.

Furthermore, the trade off property of the robustness function implies that only outcomes less desirable than the predicted outcome can have positive robustness against uncertainty. Decision analysis must, therefore, be based on exploring the robustness curve, rather than just its end-point at the predicted outcome. The robustness curves of different decision alternatives may cross one another. When this happens, the analyst may prefer one option over one range of required outcomes, while preferring the other option over a different range of outcome requirements. In other words, the analyst may encounter a reversal of preference between one option (e.g. the putative optimum) and a different option (which is putatively sub-optimal). We have seen this crossing of robustness curves and consequent potential for preference reversal in both examples.

The methodology that is developed in this paper is to satisfy a performance requirement on the outcome — rather than trying to optimize the outcome — and to maximize the robustness against uncertainty. This is a procedural optimization rather than a substantive optimization. The outcome is the substantive "good" that one seeks: small probability of large-degree cascading failures in our first example; short traffic jam duration in our second example. In the robust-satisficing approach and we only attempt to make the substantive outcome good enough. In contrast to the substantive outcome, the robustness is an aspect of the procedure of reaching a decision, and the robustness is optimized, not the outcome. Once the decision is made and implemented, the robustness is of no substantive consequence.

6 References

1. Ben-Haim, Yakov. (2006) *Info-Gap Decision Theory: Decisions Under Severe Uncertainty,* 2nd edition, Academic Press, London.

2. Ben-Haim, Yakov. (2010). *Info-Gap Economics: An Operational Introduction,* Palgrave-Macmillan, London.

3. Ben-Haim, Yakov. (2012a). Doing our best: Optimization and the management of risk, *Risk Analysis,* 32(8), 1326–1332.

4. Ben-Haim, Yakov. (2012b). Why risk analysis is difficult, and some thoughts on how to proceed, *Risk Analysis,* 32(10), 1638–1646.

5. Ben-Haim, Yakov. (2018). *The Dilemmas of* Wonderland: *Decisions in the Age of Innovation,* Oxford University Press.

6. Ben-Haim, Yakov, Zetola, Nicola M. & Dacso, Clifford. (2012). Info-gap management of public health policy for TB with HIV-prevalence, *BMC Public Health,* 12, 1091. DOI: 10.1186/1471-2458-12-1091.

7. Burgman, Mark. (2005). *Risks and Decisions for Conservation and Environmental Management,* Cambridge University Press, Cambridge.

8. Chinnappen-Rimer, S. & Hancke, G.P. (2011). Actor coordination using info-gap decision theory in wireless sensor and actor networks, *Intl. J. of Sensor Networks,* 10(4), 177–191.

9. Fraid, Oded (2016). Info-gap analysis of cascading failures in networks, MSc thesis, Technion—Israel Institute of Technology.

10. Hall, Jim W., Lempert, Robert J., Keller, Klaus, Hackbarth, Andrew, Mijere, Christophe & McInerney, David J. (2012). Robust climate policies under uncertainty: A comparison of robust decision making and info-gap methods, *Risk Analysis,* 32 (10), 1657–1672.

11. Harp, Dylan R. & Vesselinov, Velimir V. (2013). Contaminant remediation decision analysis using information gap theory, *Stochastic Environmental Research and Risk Assessment,* 27(1), 159–168.

12. Kanno, Y. & Takewaki, I. (2006). Robustness analysis of trusses with separable load and structural uncertainties, *Intl. J. of Solids and Structures,* 43(9), 2646–2669.

13. Knoke, Thomas. (2008). Mixed forests and finance — Methodological approaches, *Ecological Economics,* 65(3), 590–601.

14. Moffitt, L. Joe, Stranlund, John K. & Field, Barry C. (2005). Inspections to avert terrorism: Robustness under severe uncertainty, *J. Homeland Security and Emergency Management,* vol. 2, no. 3.

    http://www.bepress.com/jhsem/vol2/iss3/3.

15. Schwartz, Barry, Ben-Haim, Yakov & Dacso, Cliff. (2011). What Makes a Good Decision? Robust Satisficing as a Normative Standard of Rational Behaviour, *The J. for the Theory of Social Behaviour,* 41(2), 209-227.

A Degree of Cascading Failure in Triangular Hierarchical Networks, Eq.(1)

Fig. 1 shows 5 triangular hierarchical networks. In each network of this figure, the largest possible cascade of corrupted nodes is seen to be a rectangle. In the 2-row network this is a 1×2 rectangle, in the 3-row network it is a 2×2 rectangle, in the 4-row network it is a 2×3 rectangle, etc.

Thus, in a network with $2n$ rows, for $n = 1, 2, \ldots$, the maximum-failure cascade is an $n \times (n+1)$ rectangle of nodes, containing $n(n+1)$ nodes. The total number of nodes in a network with $2n$ rows is $\sum_{i=1}^{2n} i = 2n(2n+1)/2$. Thus the degree of cascading failure in a network with $2n$ rows is:

$$\phi_{2n} = \frac{n(n+1)}{2n(2n+1)/2} = \frac{n+1}{2n+1}, \quad n = 1, 2, \ldots \tag{23}$$

Similarly, in a network with $2n+1$ rows, for $n = 0, 1, \ldots$, the maximum-failure cascade is an $(n+1) \times (n+1)$ rectangle of nodes, containing $(n+1)^2$ nodes. The total number of nodes in a network with $2n+1$ rows is $\sum_{i=1}^{2n+1} i = (2n+1)(2n+2)/2 = (2n+1)(n+1)$. Thus the degree of cascading failure in a network with $2n+1$ rows is:

$$\phi_{2n+1} = \frac{(n+1)^2}{(2n+1)(n+1)} = \frac{n+1}{2n+1}, \quad n = 0, 1, 2, \ldots \tag{24}$$

which is exactly the same as eq.(23). Eqs.(23) and (24) are combined as eq.(1).

B Derivation of the Probability Distribution in Eqs.(4) and (5)

Derivation of eq.(4). A cascading failure has fractional size of zero if no nodes in the telephone chain fail. Thus $P(\psi = 0) = \pi^n$ which is the first line of eq.(4).

A cascading failure has fractional size of i/n, for $i = 1, \ldots, n$, if the first $n-i$ nodes do not fail, for which the probability is π^{n-i}, and the next node does fail for which the probability is $1-\pi$. Failure of nodes is statistically independent, so the probability that $\psi = i/n$ is the product of these two terms, which is the second line of eq.(4).

Derivation of eq.(5). The cumulative probability distribution is the sum of terms of the probability distribution:

$$\begin{aligned} P\left(\psi \le \frac{i}{n}\right) &= \sum_{j=0}^{i} P\left(\psi = \frac{i}{n}\right), \quad i = 0, 1, \ldots, n && (25) \\ &= \pi^n + \sum_{j=1}^{i} \pi^{n-j}(1-\pi) && (26) \\ &= \pi^n + (1-\pi)\pi^n \sum_{j=1}^{i} \left(\frac{1}{\pi}\right)^j && (27) \end{aligned}$$

The geometric sum in eq.(27) is:

$$\sum_{j=1}^{i} \left(\frac{1}{\pi}\right)^j = \frac{(1/\pi) - (1/\pi)^{i+1}}{1 - (1/\pi)} = \frac{1 - \pi^i}{\pi^i(1-\pi)} \tag{28}$$

Thus eq.(27) becomes:

$$\begin{aligned} P\left(\psi \le \frac{i}{n}\right) &= \pi^n + \pi^{n-i}(1-\pi^i), \quad i = 0, 1, \ldots, n && (29) \\ &= \pi^n + \pi^{n-i} - \pi^n && (30) \\ &= \pi^{n-i} && (31) \end{aligned}$$

which is eq.(5).

C Derivation of the Robustness in Eq.(10)

Let $m(h)$ denote the inner maximum in eq.(9). From eq.(6) we see that this inner maximum occurs when π is as small as possible in the uncertainty set $\mathcal{U}(h)$ of eq.(7). This occurs when $\pi = (\widetilde{\pi} - sh)^+$ where we have defined $x^+ = x$ if $x \geq 0$ and $x^+ = 0$ otherwise. Thus:

$$m(h) = 1 - \left[(\widetilde{\pi} - sh)^+\right]^{n-i} \tag{32}$$

The robustness is the greatest value of h at which this expression does not exceed $P_{\rm c}$:

$$1 - \left[(\widetilde{\pi} - sh)^+\right]^{n-i} \leq P_{\rm c} \tag{33}$$

For $h \leq \widetilde{\pi}/s$, solving this relation at equality yields the robustness in eq.(10). This expression is no greater than $\widetilde{\pi}/s$ so we needn't consider greater values of h. This completes the derivation of eq.(10).