Working Paper No. MS-4-06

The GAP in Info-Gap

Moshe Sniedovich
Department of Mathematics and Statistics
The University of Melbourne
Melbourne, Vic 3001, Australia

moshe@ms.unimelb.edu.au
www.ms.unimelb.edu.au/∼moshe

December 31, 2006

Abstract

Info-Gap is supposedly a new theory for decision making under severe uncertainty. Its claim to fame is that it is non-probabilistic in nature and thus offers an alternative to all current theories for decision making under uncertainty. In this short article I explain the Gap in Info-Gap, that is I explain how *Info-Gap* resolves the gap between the information we need and the information we have. It turns out that this is done not by a new theory but rather by a famous, more than 60-year old recipe commonly called *worst-case analysis* or simply *Maximin*.

Keywords: Decision making, severe uncertainty, maximin, worst-case analysis, info-gap.

1 Introduction

The objective of this very short essay is to clarify the GAP in *Info-Gap*, that is to explain how *Info-Gap* bridges the large gap between the information we need and the information we have in decision making under severe uncertainty.

I explained this in a rather technical language in a short paper entitled *Eureka! Info-Gap is Worst Case Analysis (maximin) is disguise!* (Sniedovich [2006b]) and in a lengthy full paper (Sniedovich [2006a].

In this paper I explain the same thing, but in a non-technical language. The point is that the issues involved are so fundamental that it is not necessary to go deep into the technical/mathematical aspects of the problem. Common sense is all that is required to understand the true nature of *Info-Gap*.

To motivate the discussion, recall that

> Info-gap decision theory is radically different from all current theories of decision under uncertainty. The difference originates in the modelling of uncertainty as an information gap rather than as a probability.
>
> Ben-Haim [2006, p.xii]

and

> An information-gap model of uncertainty is a non-probabilistic quantification of uncertainty. Info-gap models entail no measure functions: neither probability densities nor fuzzy membership functions[2]. Info-gap models concentrate on the disparity between what *is known* and what *could be known*, while making very little commitment about the structure of the uncertainty.
>
> Ben-Haim [2006, p. 2]

and

> Probability and info-gap modeling each emerged as a struggle between rival intellectual schools. Some philosophers of science have tended to evaluate the info-gap approach in terms of how it would serve physical science in place of probability. This is like asking how probability would have served scholastic demonstrative reasoning in the place of Aristotelian logic; the answer: not at all. But then, probability arose from challenges different from those which faced the scholastics, just as the info-gap decision theory which we will develop in this book aims to meet new challenges.
>
> Ben-Haim [2006, p. 12]

and

> The place to start our investigation of the difference between probability and info-gap uncertainty is with the question: can ignorance be modelled probabilistically? The answer is 'no'. The ignorance that is important to the decision maker is the disparity between *is known* and *needs to be known* in order to make a responsible decision; ignorance is an info-gap.
>
> Ben-Haim [2006, p. 12]

and so on, and so on, . . .

To cut a long story short, *Info-Gap* wants us to believe that it offers a new, radically different approach to modeling severe uncertainty in decision making, namely that it has a new recipe for bridging the gap between the information we *need* and the information we *have* in a decision making situation under severe uncertainty.

For the purposes of our discussion it is more convenient to express the same idea slightly differently: how do we make decisions in situations where the *exact value* of a parameter we use in the decision model is subject to severe uncertainty?

The standard way of dealing with generic problems of this type is to use an *estimate* of the *exact value* of the parameter of interest. That is, in the decision model we use an estimate of the exact value of the parameter of interest rather than the exact value of this parameter. We do it all the time.

Needless to say, in cases where there is *severe* uncertainty regarding the exact value of the parameter of interest, we have to be very careful the way we use estimates. After all, in this environment the estimates can be of very *poor quality*, that is, they are likely to be *substantially wrong*.

To explain how *info-Gap* deals with this fundamental issue it would be helpful to consider firstly the way **classical decision theory** copes with this fundamental, difficult issue.

2 Textbook Recipes

To appreciate the difficulty posed by severe uncertainty, it is instructive to consider less extreme situations, that is situations where the uncertainty associated with the exact value of the parameter of interest is not severe.

So recall that classical decision theory distinguishes between three types of decision making situations as far as uncertainty is concerned:

· Decision making under *certainty.*

· Decision making under *risk.*

· Decision making under *strict uncertainty.*

The first case represents situations where we pretend that there is no uncertainty at all regarding the decision making situation.

The second case represents situations where the uncertainty involved with the decision making situation can be described and quantified by conventional statistical and probabilistic models and/or methods.

The third case is the most difficult. Here we do not know much about the consequences of our decisions and therefore there is not much to work with to develop a solid, comprehensive, and useful decision making methodology.

For the purposes of our discussion it is best to regard the first two cases as “easy” and to focus on the third case.

So let us begin with the observation that classical decision theory suggests no magical foolproof recipes for dealing with such situations. What it does offer is a pair of fundamental approaches, to wit PRINCIPLES, that can be considered in situations like this.

Over the years these two principles became famous, or rather infamous, because both are very problematic. In any case, here is the celebrated duo:

· **Laplace's Principle of insufficient Reason** (1825)

· **Wald's Maximin Principle** (1945)

In brief, *Laplace's Principle* suggests that if you really do not know anything about the exact value of the parameter of interest, then it is reasonable to assume that all its "potential" values are *equally likely*.

Technically speaking, this means that you can regard the parameter of interest as a *random variable* associated with a *uniform* probability distribution function over the set of potential values this parameter can take. There are of course cases where this is impossible, eg. in cases where the set of potential values of the parameter is unbounded.

The attractiveness of this principle stems from the fact that it transforms a difficult problem (decision making under strict uncertainty) into a relatively "easy" problem (decision making under risk).

The *Maximin Principle* goes much further: it transforms a decision making situation under *strict uncertainty* into a decision making situation under *certainty*.

It does this magical trick by following my dear wife's attitude towards risk: as a rule the *worst possible thing will happen*. That is, this principle assumes that *Mother Nature* is playing *against us* in that it always selects the *least favorable* value for the parameter of interest. This, of course, is a very pessimistic view of how *Mother Nature* works, but ... this is what *worst-case analysis* is all about.

But my wife is in very a good company here. As noted by Rustem and Howe[2002], the basic idea behind worst-case analysis predates Wald (1902-1950):

> The gods to-day stand friendly, that we may,
> Lovers of peace, lead on our days to age!
> But, since the affairs of men rest still incertain,
> Let's reason with the worst that may befall.
>
> William Shakespeare (1564 - 1616)
> Julius Caesar, Act 5, Scene 1

The attractiveness of this attitude towards uncertainty is that it transforms a difficult problem (decision making under strict uncertainty) into a "very easy" problem (decision making under certainty). That is, we exploit the fact that *Mother Nature* is so antagonistic that it becomes completely predictable. This removes the uncertainty altogether and we are left with a simple deterministic problem.

2.1 Example

Suppose that one lovely morning you find four envelopes and a note on your doorstep. For your convience, the full text of the note is shown in Exhibit 1. For your convenience Table 1 depicts the information Joe provided on the four envelopes.

So what do you do Dear Sir/Madam? Which envelope should you open?

Table 2 summarizes the results obtained by applying the two principles to our little problem. Each envelope is evaluated in accordance with the two recipes. The *Wald* column selects the smallest entry in the *Possible Amounts* column, whereas the *Laplace* column computes the arithmetic average of the entries in the *Possible Amounts* column.

Good morning Sir/Madam:

I left on your doorstep four envelopes. Each contains some money. You are welcome to open any one of these envelopes and keep the money you find there.

Please note that as soon as you open an envelope the other three will automatically dissolve, so think carefully about which one of these envelopes you should open.

To help you decide what you should do, I printed on each envelope the possible amounts of money (in Australian dollars) you may find there. The amount that is actually there is equal to one of these figures.

Unfortunately the entire project is under severe uncertainty so I cannot tell more than this.

Good luck!

Joe.

Exhibit 1: Joe's Note

Envelope	*Possible Amounts (Australian dollars)*
$E1$	$20, 10, 300, 786$
$E2$	$2, 4000000, 102349, 500000000, 99999999, 56435432$
$E3$	$201, 202$
$E4$	200

Table 1: Easy Problem

Envelope	*Possible Amounts*	*Wald*	*Laplace*
$E1$	$20, 10, 300, 786$	10	279
$E2$	$2, 4000000, 10234$	2	1336745.3333 $\checkmark$
$E3$	$201, 202$	201 $\checkmark$	201.5
$E4$	200	200	200

Table 2: Results

For instance, consider the first envelope, $E1$. The super pessimistic Wald assumes that the worst value will materialize. Hence, it selects the smallest of the items on the list $20, 10, 300, 786$, which is 10.

Laplace assumes that the amount in $E1$ is a uniformly distributed random variable on this very list. Hence, the expected value is the arithmetic mean of the elements on the list: $\frac{1}{4}(20 + 10 + 300 + 786) = 279$.

In short, if you follow *Wald* you'll open the third envelope, $E3$, and if you follow *Laplace* you'll open the second envelope, $E2$.

What would you do, dear reader?

2.2 Summary

The two basic principles offered by classical decision theory for decision making under strict uncertainty transform the severe uncertainty into something easier to cope with: if we follow Laplace, we end up in a *decision making under risk* environment, if we follow Wald we end up in a *decision making under certainty* environment.

Now, since *Info-Gap* claims to be a new theory, one that is *radically different* from *all* current decision theories, we expect it to be radically different from these two principles. And since it claims to be a *probabilistic-free theory*, it is only natural to expect *Info-Gap* to explain in what way it is radically different from the *worst-case analysis* dictated by *Wald's Maximin Principle*.

Let us see.

3 Info-Gap

The first thing to note is that *Info-Gap* violates the fundamental maxim of decision making under severe uncertainty. This practical rule argues as follows:

> Fundamental Maxim of decision making under severe uncertainty
>
> Thou shalt not base your analysis on a single estimate!

That is, the centerpiece of the *Info-Gap* decision model is an estimate of the true value of the parameter of interest.

Now, of course *info-Gap* is fully aware of the fact that such an estimate is of poor quality and is likely to be substantially wrong. So to compensate for this violation of the *Fundamental Maxim*, *Info-Gap* ranks decisions on the basis of their *robustness*: the best decision is one whose robustness is the largest. That is, the objective function in the *Info-Gap* model stipulates the robustness of the decisions and the model dictates that this robustness must be maximized.

But how do you define robustness of a decision under severe uncertainty?

Well, *Info-Gap* does this by the deployment of a very simple *worst-case analysis* in the immediate neighborhood of the estimate it deploys. In other words, it insists that a certain performance requirement must be satisfied – in the worst case sense – in a region surrounding the estimate. The "size" of this region is a measure of the robustness of the decision.

In short, the best decision is one that satisfies the performance requirement – in the worst case sense – over the largest region surrounding the estimate.

This strategy is described graphically in Figure 1, where $\tilde{u}$ denotes the estimate and the regions surrounding it are circles. Each circle represents a decision: the circle representing decision a is the largest circle over which decision a satisfies the performance requirement for *all* values of u in the circle. If we attempt to expand this circle any further then some point in the expanded region will not satisfy the performance requirement. In the parlance of *worst-case analysis*, the worst point (as far as performance is concerned) in the expanded circle will not satisfy the performance requirement.

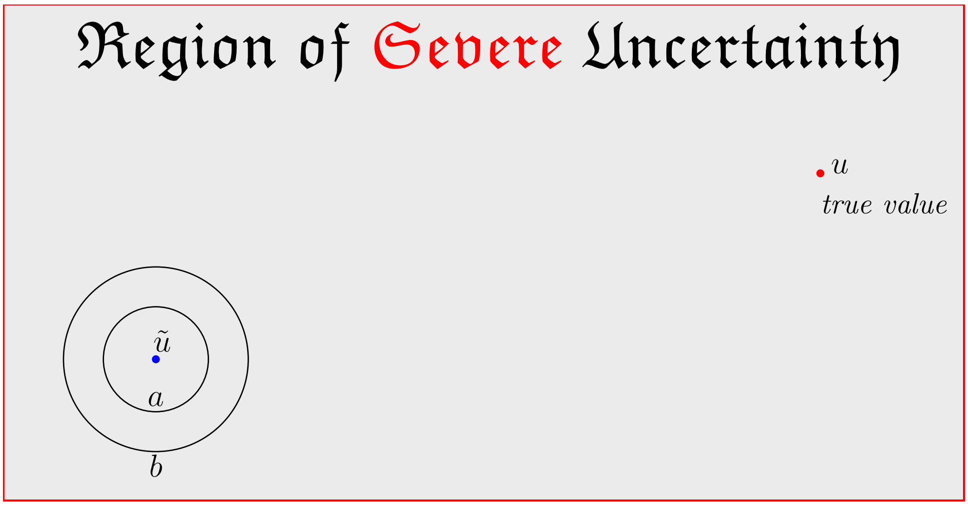


Figure 1: The fundamental flaw in *Info-Gap* treatment of severe uncertainty

Since the circle representing decision b is larger than the circle representing decision a, we regard b as a more robust, hence better, decision. If these are the only decisions available to us, we shall declare b to be the winner, hence optimal.

Needless to say, the fact that b is more robust than a in the neighborhood of the estimate $\tilde{u}$ does not imply that b is more robust than a in the neighborhood of the exact value of the parameter. So the question arises:

How does *Info-Gap* justify this very naive approach to severe uncertainty?

The answer is: it does not.

In other words, *Info-Gap* seems to be fully content with the idea that one can base decision making under severe uncertainty on an analysis of the immediate region surrounding a very poor estimate of the parameter of interest that is likely to be substantially wrong.

To put it bluntly, this does not make much sense.

In fact, this attitude reveals that *Info-Gap* suffers from a severe case of split personality. On one hand *Info-Gap* is very clear that under severe uncertainty decision making models should not be based on single point estimates because these estimates are poor and likely to be substantially wrong. But on the other hand, for some strange reason, *Info-Gap* does not heed to its own advice on this matter and happily conducts its business in clear violation of the **Fundamental Maxim**.

In any case, Figure 2 illustrates the fundamental flaw in *Info-Gap* treatment of severe uncertainty.

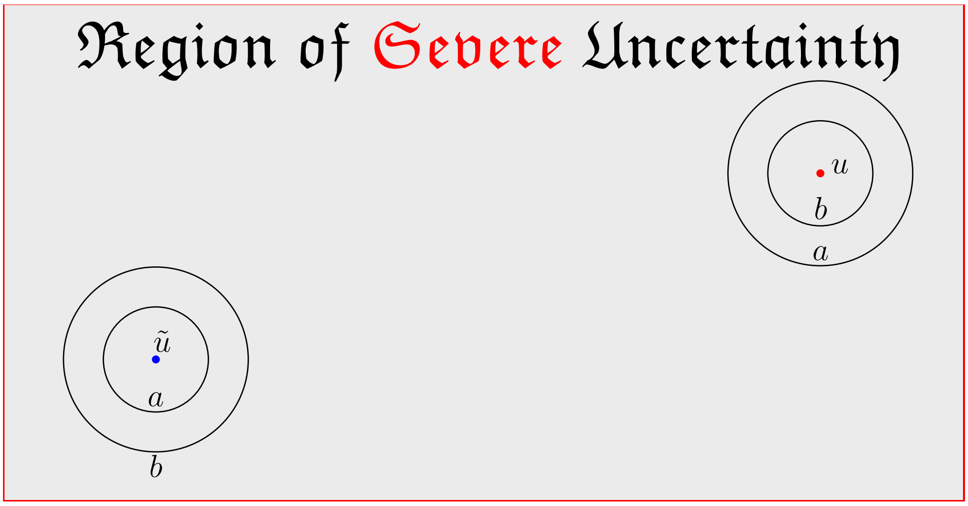


Figure 2: The fundamental flaw in *Info-Gap* treatment of severe uncertainty

It reminds us the obvious. The fact that decision a is more robust than decision b in the neighborhood around the estimate $\tilde{u}$ does not imply that a is more robust than b in the region surrounding the true value of the parameter of interest, u.

4 The GAP in *Info-Gap*

So here is the *Info-Gap* Recipe for bridging the gap between the the information we need (true value of u) and the information we have (a poor estimate of u) in a probabilistic-free style:

- Use whatever estimate you have, knowing full well that it is of poor quality and can be substantially wrong.
- Conduct *worst-case analysis* in regions surrounding the poor estimate.

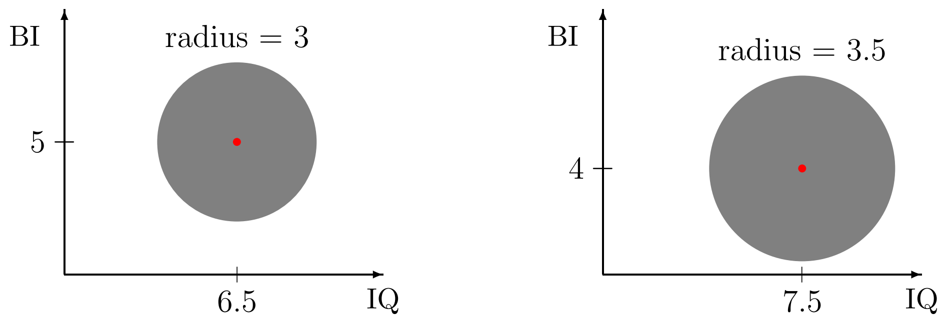
In short, *Info-Gap* is a specialized *worst-case analysis* conducted around a poor estimate of the parameter of interest.

The funny thing is that *Info-Gap* is apparently unaware that this is actually what it does. How would you explain *Info-Gap*'s claim that in the framework of its uncertainty region there is no worst case?

5 Illustrative Example

You plan to buy a present for you dog Rex, a beautiful 7 year old German shepherd. This year you decided to buy him an educational game. There are two brands in your local pet shop, **Charisma** and **Agility**. The manuals of these games provide the operating charts shown in Figure 3. The games are suitable only for dogs whose BI and IQ scores are within the shaded areas on the charts[1].

The question is: which brand should you buy for Rex – **Charisma** or **Agility**?



Recommended operating region **Charisma** | Recommended operating region **Agility**

Figure 3: Operating Charts for the two brands

We shall now briefly discuss four versions of this problem. Three are associated with different levels of *uncertainty* pertaining to Rex's BI and IQ scores, and one has nothing to do with uncertainty.

Observe that in the framework of the *Info-Gap* model, here we have two decisions (brands) and the operating charts are the regions over which the respective decisions (brands) satisfy the performance requirements. The parameter of interest u is the pair of (IQ, BI) scores.

Version 1: Certainty.

Here we assume that we have the exact scores for Rex. The choice seems to be obvious: we can choose any brand as long as Rex's scores are within the specified operating region of the brand.

Version 2: Strict Uncertainty.

Suppose that we do not have any information about Rex's scores. Given the extreme level of uncertainty, it seems that the best thing to do is go for the brand whose operating region is the largest. In our case this is the operating region of **Agility**, so it looks like this would be the best choice.

Note that to apply *Info-Gap* here we would need an estimate of the true value of the scores.

Version 3: Pretty good estimates.

Suppose that we do not have the exact values of Rex's scores, but we do have pretty good estimates of these scores. Let $\underline{a}$ denote the estimate of the IQ score and

[1] BI is short for *Barking Index.*

let $\underline{b}$ denote the estimate of the BI score. In this case we may wish to *play it safe* and select the brand providing the *largest* SAFE *deviation* from the estimates.

We shall conduct the *worst-case analysis* graphically. You are encouraged to do it analytically on your own.

Suppose that the two estimates are $\tilde{u} = (\underline{a}, \underline{b}) = (6, 6)$. How far can we go from these estimates and still be in the operating region of a brand? To answer this question, we can draw circles centered at the point $\tilde{u} = (6, 6)$ on the charts. We increase the radius of these circles until they are not FULLY contained in the operating regions, as shown in Figure 4.

Clearly, for these estimates **Charisma** seems to be far safer as the radius of the largest safe circle on its chart is much larger than the radius of the largest safe circle on the **Agility** chart.

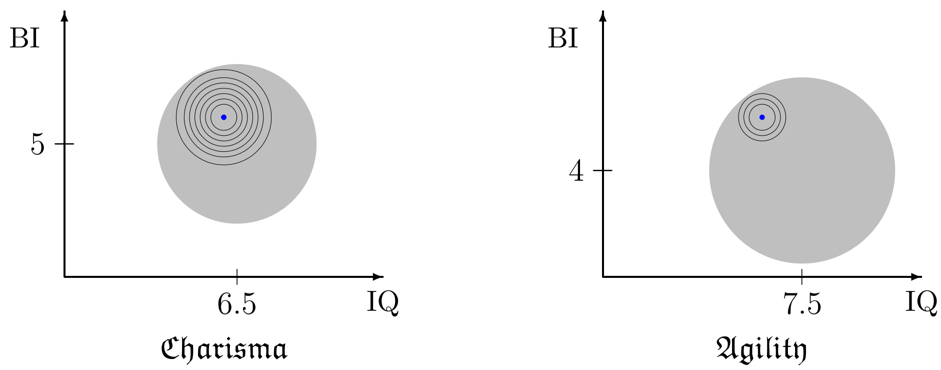


Figure 4: Worst Case Analysis

It is important to note that if the estimates we have are not good, the analysis could be much more complicated. For example, suppose that the estimates are poor and all we know is that the true values are somewhere on the line segment connecting the two end points $(6, 6)$ and $(9, 5)$ on the IQ/BI plane.

Figure 5 displays the *worst-case analysis* for these two end points. Note that **Charisma** seems to be the better brand for the point $(6, 6)$ whereas **Agility** seems to be the better brand for the point $(9, 5)$.

So, which brand should you buy for Rex in this case?

The point is that if you do not have a good estimate and you try to use a number of estimates, it is not clear which one of them should be used to determine the best decision.

Interestingly, *Info-Gap* behaves as if the estimate it uses is of very good quality so there is no need to consider other estimates. But how can you get a good estimate under severe uncertainty? You can't.

The next version we examine has nothing to do with uncertainty. I discuss it to emphasize that the scope of *worst-case analysis* goes beyond problems dealing with uncertainty.

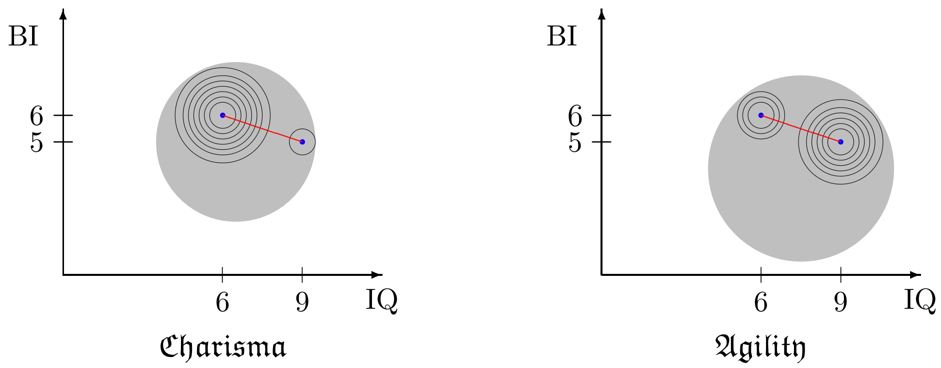


Figure 5: Worst Case Analysis

Version 4: Variability

Consider the case where before you went shopping with Rex, you called Jack the Vet and were told that Rex's scores are precisely 7 for the IQ and 7 for BI.

What do you do in this case?

Since both brands are suitable for Rex in this case, it really does not matter which brand you decide to buy. But, on second thought, how about … Rex's friends? They always play with Rex's toys and games.

In view of this additional consideration, you now want a brand that will be suitable not only for Rex, but also for his friends. In short, you want a brand that will be suitable for Rex but capable of handling the largest possible variability from Rex's scores. Since Rex's friends have similar IQ and BI scores, you decided to conduct the *worst-case analysis* in the *immediate neighborhood* of Rex's scores on the IQ/BI plane.

So formally, you are interested in the brand whose operation chart can cope (safely) with the largest deviation from the point $(7,7)$ on the IQ/BI plane.

The solution generated by the *worst-case analysis* for this version of the problem is shown in Figure 6. The clear winner is no doubt **Charisma**, so it looks like Rex and his friends will play **Charisma** for the rest of the year.

Hopefully, as promised by the manufactures, this will increase their IQ scores and decrease their BI scores.[2]

6 Conclusions

The basic strategy deployed by *Info-Gap* to bridge the information gap between the the information we need and the information we have violates the **Fundamental Maxim** of decision making under severe uncertainty.

The idea to use *worse case analysis* to transform a probabilistic problem into a deterministic problem is more than 60 years old and is well known in the decision theory literature as *Wald's Maximin Principle.*

[2]In subsequent papers on this subject I'll report on Rex's progress on the IQ/BI front.

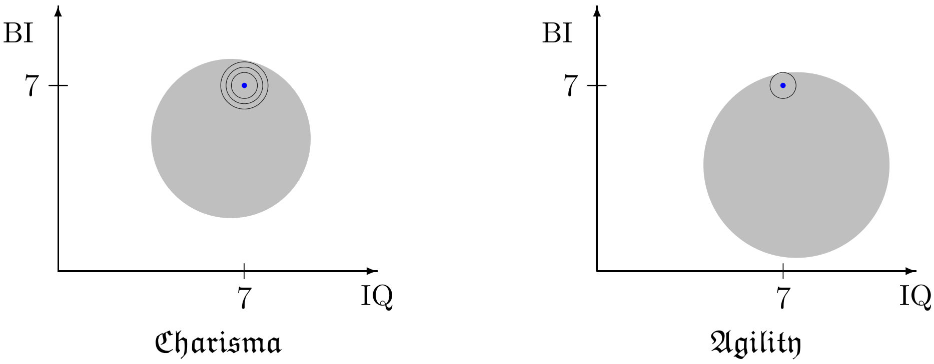


Figure 6: Worst Case Analysis

It is not clear therefore what *Info-Gap* contributes to the state of the art in decision making under severe uncertainty.

Technical details concerning the role and place of *Info-Gap* in decision theory can be found in Sniedovich [2006a, 2006b].

References

[1] Ben-Haim Y, *Info-Gap Decision Theory*, Elsevier, 2006.

[2] Rustem B, Howe M, *Algorithms for Worst-case Design and Applications to Risk Management*, Princeton University Press, 2002.

[3] Sniedovich M, What is Wrong with Info-Gap? An Operations Research Perspective, Paper No. MS-1-06, 2006a.

[4] Sniedovich M, Eureka! Info-Gap is Worst Case Analysis (maximin) in Disguise!, Paper No. MS-2-06, 2006b.

[5] Sniedovich M, The Fundamental Flaw in Info-Gap, Paper No. MS-3-06, 2006c.