Lecture Notes on Value Judgements

Yakov Ben-Haim
Yitzhak Moda'i Chair in Technology and Economics
Faculty of Mechanical Engineering
Technion — Israel Institute of Technology
Haifa 32000 Israel
yakov@technion.ac.il
https://yakovbh.net.technion.ac.il/

Source material: Yakov Ben-Haim, 2006, *Info-Gap Decision Theory: Decisions Under Severe Uncertainty,* 2nd edition, Academic Press. Chapter 4.
A Note to the Student: These lecture notes are not a substitute for the thorough study of books. These notes are no more than an aid in following the lectures.

"It is true that convictions can best be supported with experience and clear thinking. On this point one must agree with the extreme rationalist. The weak point of his conception is, however, this, that those convictions which are necessary and determinant for our conduct and judgments, cannot be found solely along this solid scientific way.

"For the scientific method can teach us nothing else beyond how facts are related to, and conditioned by, each other. ... [Yet] knowledge of what 'is' does not open the door directly to what 'should be'. One can have the clearest and most complete knowledge of what 'is', and yet not be able to deduct from that what should be the 'goal' of our human aspirations." Albert Einstein, *Out Of My Later Years,* pp21–22. Thames and Hudson Pub., London, 1950.

Contents

[0]lectures\risk\lectures\vjud02.tex 30.12.2020

1 The Problem

¶ We have learned that, for $\widehat{h}(q, r_c)$, **bigger is better.**
But now we face the questions:
- How big is big enough?
- What is the utility of a marginal increment in $\widehat{h}$?
- How do we evaluate the **numerical value** of $\widehat{h}$ in terms of

subjective personal or **social values?**

¶ A similar problem arises in probabilistic decision theory.
If P_s = probability of success,
Then **bigger is better** for P_s just as for $\widehat{h}$.
But the same value judgment arises about P_s as about $\widehat{h}$:
How big is big enough?

¶ Value judgment as calibration:
How to **calibrate** $\widehat{h}$ in
subjective, value-based linguistic terms
such as: big enough, too small, etc.

¶ We will consider three methods for calibrating $\widehat{h}$:
1. Prior experience.
2. Dimensional normalization.
3. Severity of Consequences.

All three are in fact cases of **reasoning by analogy.**

¶ **Calibration: incremental or absolute?**
- Incremental calibration: how significant is $\Delta\widehat{h}$?
- Absolute calibration: how robust (i.e. safe) is $\widehat{h}$?

- Incremental calibration: plausible answer with analogical reasoning.
- Absolute calibration:
  - ○ No solution.
  - ○ Moral version of absolute space: does not have meaning.

2 Prior Experience

¶ Suppose we have a system S_1 which:
1. We are familiar with.
2. Performs well.
3. Has robustness $\widehat{h}_1$.

¶ Now consider a system S_2 which:
1′. Is similar to S_1 in important respects.
2′. Has unknown performance.
3′. Has robustness $\widehat{h}_2$ which is close to $\widehat{h}_1$.

¶ We may tend to conclude that,
by analogy between S_1 and S_2,
S_2 will tend to perform well, like S_1.
This suggests that $\widehat{h}_2$ is adequate robustness.

¶ Our reasoning here is:
1. S_1 and S_2 are similar.
2. S_1 performs well.
3. Therefore S_2 will tend to perform well;
   and $\widehat{h}_2$ will tend to be adequate.

This is **reasoning by analogy:**
Things which are **similar in some respects**
will **tend to be similar in other respects** as well.
Reasoning by analogy is **not proof.**
Reasoning by analogy is **plausible inference.**

3 Dimensional Normalization

¶ The horizon of uncertainty, h, of an info-gap model may have **physical units.**
For instance, if $u(x)$ = uncertain quantity of pollutant, [kg/m^2],
whose info-gap model is:

$$\mathcal{U}(h,\widetilde{u}) = \{u(x) : |u(x) - \widetilde{u}(x)| \le h\psi(x)\}, \quad h \ge 0 \tag{1}$$

we see that $[h] = [u(x)] = [\widetilde{u}(x)] = [\text{kg/m}^2]$. ($\psi(x)$ is a dimensionless envelope.)

Or, for example, if $u(t)$ = uncertain mechanical stress, [N/m^2], whose info-gap model is:

$$\mathcal{U}(h,\widetilde{u}) = \left\{u(t) : \int_0^\infty (u(t) - \widetilde{u}(t))^2 \, \mathrm{d}t \le h^2\right\}, \quad h \ge 0 \tag{2}$$

then $[h] = [u]\sqrt{\text{sec}} = \dfrac{\text{N}\sqrt{\text{s}}}{\text{m}^2}$.

¶ The robustness $\widehat{h}$ is the least upper bound of a set of h-values:

$$\widehat{h} = \max\{h : \text{ Some critical condition holds}\} \tag{3}$$

Thus $\widehat{h}$ and h always have the same units.

¶ This implies that:
- There is **no inherent calibration** of $\widehat{h}$ in absolute units.
- We cannot say '$\widehat{h} = 10^6$' implies '$\widehat{h}$ is large'.
- We must calibrate the units of $\widehat{h}$ in some way.

¶ One type of calibration of $\widehat{h}$ is based on:
- physical reasoning,
- engineering judgment,
- economic judgment,
- etc.

These considerations can **normalize** $\widehat{h}$ to **dimensionless** form.

¶ For instance, consider the envelope-bound info-gap model in eq.(1) on p. 4.
The robustness is dimensional: $[\widehat{h}] = [\widetilde{u}] = \text{kg/m}^2$.
However, $\dfrac{\widehat{h}}{\widetilde{u}}$ is dimensionless.

$\dfrac{\widehat{h}}{\widetilde{u}} \gg 1$ suggests: Fluctuations which are large
w.r.t. the nominal function
do not cause failure.
suggests: plausibly reliable system.

while

$\dfrac{\widehat{h}}{\widetilde{u}} \ll 1$ suggests: Fluctuations which are small
w.r.t. the nominal function
entail the possibility of failure.
suggests: plausibly risky system.

¶ For instance, consider the energy-bound info-gap model in eq.(2) on p. 4.
The robustness is dimensional: $[h] = [u]\sqrt{\text{sec}} = \frac{\text{N}\sqrt{\text{s}}}{\text{m}^2}$.
Suppose T is a typical time constant of the system. E.g.:
- Decay time.
- Transient duration.
- Lifetime.
- Etc.

$\widetilde{u}(t)$ is a typical u-function.
So $\widetilde{u}(t)\sqrt{T}$ is a typical-valued function
whose dimensions are the same as $\widehat{h}$:
$[\widehat{h}] = [\widetilde{u}(t)\sqrt{T}]$.

Thus $\frac{\widehat{h}}{\widetilde{u}(t)\sqrt{T}}$ is dimensionless.

If $\frac{\widehat{h}}{\widetilde{u}(t)\sqrt{T}} \gg 1$ then all functions "much larger" than typical do not cause failure.
$\widehat{h}$ is plausibly large enough;
The system is plausibly reliable.

If $\frac{\widehat{h}}{\widetilde{u}(t)\sqrt{T}} \ll 1$ then there are functions "close" to typical which cause failure.
$\widehat{h}$ is plausibly too small;
The system is plausibly risky.

¶ Calibration by **dimensional normalization**
is a special case of **reasoning by analogy.**

Recall that an analogical inference is:
Things which are similar in some respects
will tend to be similar in other respects.

¶ For example, when we concluded that:

$\frac{\widehat{h}}{\widetilde{u}(t)\sqrt{T}} \gg 1$ implies 'high robustness'

what analogy are we making?

1. Since $\frac{\widehat{h}}{\widetilde{u}(t)\sqrt{T}} \gg 1$
   Only loads which are:
   very different from $\widetilde{u}(t)$
   cause failure.
2. Such loads seem rare, extraordinary, unusual.
3. Thus, by analogy, failure seems rare, extraordinary, unusual
   and the system seems reliable.

4 Analogical Reasoning

¶ Now we consider more rigorously the problem of **calibrating** $\widehat{h}(q, r_c)$.

¶ Stating the problem more quantitatively:
$\widehat{h}(q, r_c)$ depends on the decision vector q.
Let q_1 and q_2 be two different decisions, for which:
$\widehat{h}(q_1, r_c) > \widehat{h}(q_2, r_c)$.
This means that, in terms of robustness:
q_1 is **better than** q_2.
But by how much?
Is q_1 substantially, **qualitatively better** than q_2?
We seek an evaluation in **linguistic terms**
of **quantitative** increments in $\widehat{h}(q, r_c)$.

¶ **Analogy, metaphor and translation:**

- We want to translate $\widehat{h}$ or $\Delta\widehat{h}$ from math to natural human language.
- We want to extend the meaning of $\widehat{h}$ or $\Delta\widehat{h}$.
- Metaphor extends meaning. E.g. "river of time".
- Analogical reasoning is metaphor between two languages.

¶ For any two entities, f_1 and f_2,
If f_1 is **"significantly greater"** than f_2
we will write: $f_1 \succ f_2$.
These entities may be **numbers**
or **words:** linguistic values.
Linguistic example: $f =$ reliability of a system.
where f takes the values "low", "moderate", "high", etc.
Then "high" $\succ$ "low".

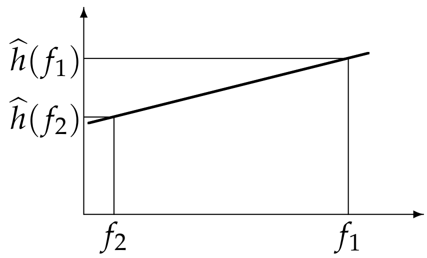

Figure 1: Illustration of calibration by analogy.

¶ Now suppose that the robustness function $\widehat{h}(f)$ increases with f (see fig. 1):

$$f_1 \succ f_2 \tag{4}$$

and:

$$\widehat{h}(f_1) > \widehat{h}(f_2) \tag{5}$$

Eqs.(4) and (5) **analogically imply** that:

$$\widehat{h}(f_1) \succ \widehat{h}(f_2) \tag{6}$$

¶ Eqs.(4)–(6) are an **inference by analogy:**
1. The f values are **ranked** by significant increments.
2. The f values and the $\widehat{h}$ values agree in **ordinal ranking.**
3. Hence the $\widehat{h}$ values **tend to be ranked** by significant increments.

¶ Recall that an **analogical inference** is:
Things that **are** similar in **some** respects,
will **tend** to be similar in **other** respects.
Analogy is **not proof.**
Analogy is **plausible explanation.**

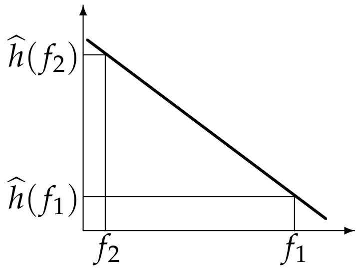

Figure 2: Illustration of calibration by reverse analogy.

¶ There is also reasoning by **reverse analogy.** (see fig. 2)
Instead of: (4) and (5) analogically imply (6).
We have:

$$f_1 \succ f_2 \tag{7}$$

and

$$\widehat{h}(f_2) > \widehat{h}(f_1) \tag{8}$$

analogically imply that:

$$\widehat{h}(f_2) \succ \widehat{h}(f_1) \tag{9}$$

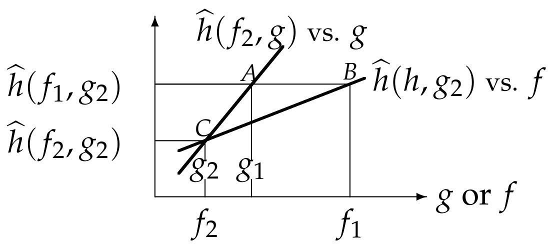

Figure 3: Illustration of calibration by analogy with two variables.

¶ We can also **concatenate** analogical inferences (fig. 3).
Suppose that $\widehat{h}$ depends on two variables: $\widehat{h}(f, g)$.

1. First analogical inference:

$$f_1 \succ f_2 \quad \text{and} \quad \widehat{h}(f_1, g_2) > \widehat{h}(f_2, g_2) \tag{10}$$

analogically imply that:

$$\underbrace{\widehat{h}(f_1, g_2)}_{B} \succ \underbrace{\widehat{h}(f_2, g_2)}_{C} \tag{11}$$

2. Suppose $\widehat{h}(f_2, g)$ increases with g.
Suppose that g_1 is a value such that:

$$\underbrace{\widehat{h}(f_2, g_1)}_{A} = \underbrace{\widehat{h}(f_1, g_2)}_{B} \tag{12}$$

3. Second analogical inference: Inference (11) ($B \succ C$) with:

$$\underbrace{\widehat{h}(f_2, g_1)}_{A} = \underbrace{\widehat{h}(f_1, g_2)}_{B} \quad \text{and} \quad g_1 > g_2 \tag{13}$$

analogically imply that:

$$g_1 \succ g_2 \tag{14}$$

¶ We now consider an example of analogical inference for calibrating $\widehat{h}$ by **severity of consequences.**

5 Calibration by Consequence Severity

¶ Surface treatment problem:
- Determine program of treatment for a region.
  E.g. apply protective material, apply fertilizer, allow grazing, etc.
- Model: degree of influence as function of degree of treatment on the region.
- Use model to determine degree of treatment.
- Model is uncertain.
- Determine the treatment to satisfice the level of influence.
- Evaluate robustness function.
- Numerical values of influence are understood linguistically.
- Calibrate robustness function to find significant increments of $\widehat{h}$.
- Use calibrated robustness function to determine degree of treatment.

¶ First we evaluate $\widehat{h}(q, r_c)$.
Then we use analogical reasoning by consequence severity
to calibrate $\widehat{h}(q, r_c)$ and determine treatment.

5.1 Robustness function

¶ Basic terms:
x = position in region. $0 \le x \le 1$.
$q(x)$ = density of treatment at x.
$b(x)$ = uncertain influence function.
$R(q, b)$ = benefit to region due to treatment. Numerical values understood qualitatively.

¶ Model of benefit function:

$$R(q, b) = \int_0^1 b(x)q(x)\,\mathrm{d}x \tag{15}$$

$b(x)$ = uncertain influence function.
= benefit at x from unit treatment at x.
Note: the **model** $R(q, b)$ for how treatment q influences the region is **uncertain.**

¶ What we know about the uncertain part of the model, $b(x)$:
- The nominal function, $\widetilde{b}(x)$.
- $b(x)$ deviates from $\widetilde{b}(x)$:
  - ○ The deviation is not erratic.
  - ○ Sudden shifts in $u(x)$ are not plausible.
- Variation of $b(x)$ is greatest in the mid-region.
- Variation of $b(x)$ vanishes at the boundaries.

¶ Use an envelope-slope-bound info-gap model:

$$\mathcal{U}(h,\widetilde{b}) = \left\{ b(x) = \widetilde{b}(x) + u(x) : \ u(0) = u(1) = 0, \ \left| \frac{\mathrm{d}u(x)}{\mathrm{d}x} \right| \le h\psi(x) \right\}, \quad h \ge 0 \qquad (16)$$

where $\psi(x)$ is the known envelope function.

¶ Critical performance requirement:

$$R(q,b) \ge r_{\mathrm{c}} \qquad (17)$$

¶ Robustness function:

$$\widehat{h}(q,r_{\mathrm{c}}) = \max \left\{ h : \ \left(\min_{b \in \mathcal{U}(h,\widetilde{b})} R(q,b) \right) \ge r_{\mathrm{c}} \right\} \qquad (18)$$

¶ The benefit function, eq.(15) on p.10, can be written:

$$R(q,b) = \underbrace{\int_0^1 \widetilde{b}(x) q(x) \, \mathrm{d}x}_{\widetilde{R}(q)} + \int_0^1 u(x) q(x) \, \mathrm{d}x \qquad (19)$$

$\widetilde{R}(q)$ is the nominal, error-free benefit from treatment $q(x)$.
¶ Evaluate the robustness function:

Define the following known function:

$$\eta(x) = \int_0^x q(x') \, \mathrm{d}x' \quad \Longrightarrow \quad \mathrm{d}\eta(x) = q(x)\mathrm{d}x \qquad (20)$$

The right-most integral in eq.(19) is found by integration by parts $\left(\int_a^b g \mathrm{d}f = gf|_a^b - \int_a^b f \mathrm{d}g \right)$:

$$\int_0^1 u(x) q(x) \, \mathrm{d}x = \int_0^1 u(x) \, \mathrm{d}\eta(x) \qquad (21)$$

$$= \underbrace{u(x)\eta(x)\Big|_0^1}_{0} - \int_0^1 \eta(x) \frac{\mathrm{d}u(x)}{\mathrm{d}x} \, \mathrm{d}x \qquad (22)$$

Hence the minimum benefit up to info-gap h is:

$$\min_{b \in \mathcal{U}(h,\widetilde{b})} R(q,b) = \widetilde{R}(q) - h \int_0^1 |\eta(x)| \, \psi(x) \, \mathrm{d}x \qquad (23)$$

Equating this to r_{c} and solving for h yields $\widehat{h}$:

$$\widehat{h}(q,r_{\mathrm{c}}) = \frac{\widetilde{R}(q) - r_{\mathrm{c}}}{\int_0^1 |\eta(x)| \, \psi(x) \, \mathrm{d}x} \qquad (24)$$

unless this is negative, in which case $\widehat{h}(q,r_{\mathrm{c}}) = 0$.

5.2 Special Case

¶ Assumptions:
q = density of treatment = constant over $x \in [0,\ 1]$.
$\widetilde{b}(x) = b_0 + b_1 \sin \pi x$ = known nominal influence function.
$\psi(x)$ = known envelope function = $\sin \pi x$.
¶ With these assumptions the robustness function becomes:

$$\widehat{h}(q, r_c) = \pi b_0 + 2b_1 - \frac{\pi r_c}{q} \tag{25}$$

Eq.(25) shows that, at fixed r_c, $\widehat{h}(q, r_c)$ **increases** with increasing q, fig. 4.
Eq.(25) shows that, at fixed q, $\widehat{h}(q, r_c)$ **decreases** with increasing r_c, fig. 4.

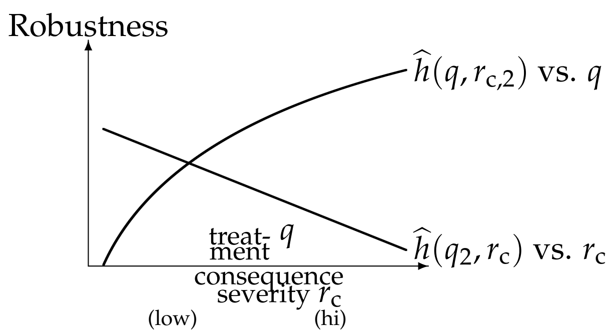

Figure 4: Robustness $\widehat{h}(q, r_c)$ vs. r_c and vs. q, eq.(25).

¶ Questions:
- What is a significant increment in $\widehat{h}$?
- What is a significant increment in q?

5.3 Calibration by Consequence Severity

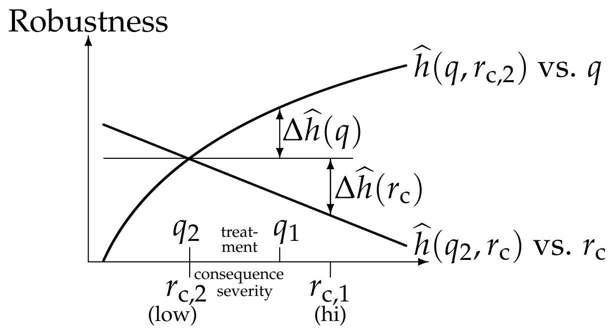

Figure 5: Robustness $\widehat{h}(q, r_c)$ vs. r_c and vs. q, eq.(25).

¶ We answer these questions, for the special case, by consequence-severity calibration.

¶ Fig. 5 shows two robustness curves.
• The decreasing straight line shows trade-off of $\widehat{h}$ vs. r_c.
• The increasing curve shows augmented $\widehat{h}$ resulting from increased treatment q.
• Two values of reward (benefit) are shown: $r_{c,1} > r_{c,2}$.
• These levels of reward are understood intuitively:
$r_{c,1}$ is substantially greater benefit than $r_{c,2}$:
$r_{c,1} \succ r_{c,2}$.
$r_{c,1}$ is "high" benefit.
$r_{c,2}$ is "low" benefit.

¶ Now concatenate 2 analogical inferences in fig. 5.
- • First inference: fixed treatment q_2:
  $\Delta\widehat{h}(r_c)$ = increment in robustness from $r_{c,1}$ to $r_{c,2}$.
  By reverse analogical inference:
  $\Delta\widehat{h}(r_c)$ = significant increment in robustness.
  That is:
  $r_{c,1} \succ r_{c,2}$ and $\widehat{h}(q_2, r_{c,2}) > \widehat{h}(q_2, r_{c,1})$ implies
  $\widehat{h}(q_2, r_{c,2}) \succ \widehat{h}(q_2, r_{c,1})$.
  That is: $\Delta\widehat{h}(r_c)$ is "large".
- • 2nd inference.
  $\Delta\widehat{h}(q)$ = increment in robustness from q_2 to q_1 at fixed benefit $r_{c,2}$.
  Choose q_1 so that $\Delta\widehat{h}(q) = \Delta\widehat{h}(r_c)$
  Now, with 1st inference ($\Delta\widehat{h}(r_c)$ is "large"): we conclude:
  $\widehat{h}(q_1, r_{c,2}) \succ \widehat{h}(q_2, r_{c,2})$.
  Thus the increment in treatment from q_2 to q_1
  is significant in analogy to the corresponding increment in $\widehat{h}$.

¶ Summary of this analogical argument:

1. The increment $\Delta h(q)$ from q_2 to q_1 (at fixed r_c) agrees in robustness to the increment $\Delta h(r_c)$ from $r_{c,1}$ to $r_{c,2}$.
2. The increment $\Delta h(r_c)$ from $r_{c,1}$ to $r_{c,2}$ is a qualitatively significant change in benefit, by analogy to $r_{c,2} \succ r_{c,1}$.
3. Thus the increment in treatment from q_2 to q_1 is expected, by analogy, to be significant.