How to keep an expert honest
or
how to prevent misinformation

Flemming Topsøe

Department of Mathematical Sciences
University of Copenhagen

Seminar, Copenhagen, September 2. 2009

CONTENT

PART I: how to prevent misinformation, abstractly and in concrete probabilistic models. Entropy and divergence

PART II: Philosophy connecting truth, belief and knowledge with applications to probabilistic models related to Tsallis entropy.

PART III: Outlook

. . . .

MORE INFORMATION: See
http://www.math.ku.dk/ topsoe/manuscripts/
(The ms. "Paradigms of Cognition" will be available primo October, I think)

Adviser $\leftrightarrow$ Investor

x: Advisers best advice.
y: Advice actually given.

For dubious reasons, misinformation ($y \neq x$) may take place.

task: prevent misinformation!

$\cdots\cdots$

Same problem in another setting: Expert$\leftrightarrow$ You , say a meteorologist giving advice to you.
task: keep expert honest, encourage professionalism!

Solution strategy

Select a function $\Phi = \Phi(x, y)$ such that

$$\Phi(x, y) \geq \Phi(x, x)$$

with equality only if there is a perfect match, i.e. $y = x$.

Strategy: Pay Adviser an up-front fee and supply with an insurance scheme, according to which Adviser has to pay you a penalty $\Phi(x, y)$ as soon as you have seen what really happened.

This prevents misinformation, encourages honesty!

The principle (penalty minimized only if there is a perfect match) we call the Perfect Match Principle (PMP) and a function Φ with the required properties we call a PMP-function.

entropy and divergence (assuming Φ is PMP-function)

Define:
entropy $=$ minimal penalty and divergence $=$ excess penalty:

$$H(x) = \Phi(x, x)\,,$$
$$D(x, y) = \Phi(x, y) - \Phi(x, x)\,.$$

Then $\Phi(x, y) = H(x) + D(x, y)\,,$ the linking identity.

The PMP-property may be turned into the inequality

$$D(x, y) \geq 0 \text{ with equality iff } y = x\,.$$

This is the fundamental inequality (FI).

... but how to find a PMP-function? Let us turn to an example!

Example: A probabilistic model

$\mathbb{A}$	x	y	local penalty
.	.	.	.
i	x_i	y_i	$\phi(x_i, y_i)$
.	.	.	.

x's and y's are *probability distributions* over an alphabet $\mathbb{A}$. Assume that $\Phi(x, y) = \sum_{i \in \mathbb{A}} \phi(x_i, y_i)$ where $\phi(s, t) = s\kappa(t)$ for a suitable penalty generator κ (a smooth function on $[0, 1]$ with $\kappa(1) = 0$, normalized so that $\kappa'(1) = -1$).

$$\Phi \text{ is average penalty: } \Phi(x, y) = \sum_{i \in \mathbb{A}} x_i \kappa(y_i)\,.$$

Theorem: Only *one* penalty generator turns Φ into a PMP-function: $\kappa(t) = \ln \frac{1}{t}$. With this choice, entropy becomes Shannon entropy and divergence becomes Kulback-Leibler divergence: $H(x) = \sum x_i \ln \frac{1}{x_i}$, $D(x, y) = \sum x_i \ln \frac{x_i}{y_i}$.

the pointwise fundamental inequality

Proof of PMP-property when $\kappa(t) = \ln \frac{1}{t}$:

$$\text{First adjust } \phi\text{: } \tilde{\phi}(s,t) = \phi(s,t) + t\,,$$

then, instead of proving

$$\sum \phi(x_i, y_i) \geq \sum \phi(x_i, x_i)\,,$$

we prove that

$$\sum \tilde{\phi}(x_i, y_i) \geq \sum \tilde{\phi}(x_i, x_i)\,.$$

Even the pointwise fundamental inequality (PFI) holds:

$$\tilde{\phi}(s,t) \geq \tilde{\phi}(s,s)\,.$$

Indeed,

$$s \ln \frac{1}{t} + t \geq s \ln \frac{1}{s} + s$$

as

$$s \ln \frac{s}{t} \geq s\left(1 - \frac{t}{s}\right) = s - t\,.$$

□

PART II: philosophy ...

truth, belief and knowledge

Before, x was a kind of truth. Now change the underlying interpretations so that x is truth, conceived as “genuine truth” or even “absolute truth”– a truth which is held by Nature .

Introduce also Observer, (could be a physicist, a statistician or a probabilist).

Nature does not give advice. So instead of advice, we now take y to represent Observers belief.

Knowledge is conceived as the *synthesis of extensive experience*, in probabilistic models corresponding to the regime where the laws of large numbers have taken over.

Worlds and interactors

The interaction Nature$\leftrightarrow$Observer takes place in a world, $\mathcal{W}$. Observer studies situations from the world, characterized by instances of the central concepts.

Another aspect of knowledge: *it is the way the world presents truth to Observer in any situation*.

C r u c i a l assumption: This depends on the truth instance x *and* on the belief instance y.

The knowledge instance, z, is thus a function of x and y:

$$z = \Pi(x, y) \text{ with } \Pi \text{ the interactor.}$$

examples of worlds

The classical world, $\mathcal{W}_1$, is characterized by the interactor

$$\Pi_1(x, y) = x\,.$$

It is a *world of observable truth*. *What you see is what is true* (WYSIWIT).

Another extreme: a black hole, $\mathcal{W}_0$, characterized by the interactor

$$\Pi_0(x, y) = y\,.$$

In this world, *what you see is what you believe* (WYSIWYB).

Mixtures: $\Pi_q(x, y) = qx + (1 - q)y$ define Tsallis worlds,$\mathcal{W}_q$.

All worlds we shall consider will be sound, that is $\Pi(x, y) = x$ provided *belief matches truth* ($y = x$).

probabilistic modeling

$\mathbb{A}$	x	y	z	effort
$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$
i	x_i	y_i	z_i	$\phi(x_i, y_i) = z_i \kappa(y_i)$
$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$

Modeling as before except now

$$\Pi(x, y) = \big(\pi(x_i, y_i)\big)_{i \in \mathbb{A}}$$

with π the local interactor and

$$\phi(s, t) = \pi(s, t)\kappa(t)$$

Interpretations: Now Φ determines the description effort or the description cost and κ is the generator. The PMP is now the natural principle for Observer to strive for a generator which ensures that effort is minimized when there is a perfect match.

Theorem: Given π, only one generator could possibly lead to a Φ which satisfies the PMP.

proof

Proof: Assume PMP holds with κ as generator. For $0 < t < 1$ put

$$f(t) = \chi(t)\kappa(t) + t\kappa'(t)$$

where

$$\chi(t) = \frac{\partial \pi}{\partial t}(t, t)\,.$$

Consider a probability vector $x = (x_1, x_2, x_3)$ with positive point probabilities. By PMP F given by

$$F(y) = F(y_1, y_2, y_3) = \sum_{1}^{3} \pi(x_i, y_i)\kappa(y_i)$$

on $]0,1[\times]0,1[\times]0,1[$ assumes its minimal value for the interior point $y = x$ when restricted to probability distributions. Thus, there exists a Lagrange multiplier λ such that

... continued ...

$$\frac{\partial}{\partial y_i}\Big(F(y_1, y_2, y_3) - \lambda \sum_1^3 y_i\Big) = 0 \text{ for } i = 1, 2, 3$$

when $y = x$. This shows that $f(x_1) = f(x_2) = f(x_3)$. Using this with $(x_1, x_2, x_3) = (\frac{1}{2}, x, \frac{1}{2} - x)$ for $x \in]0, \frac{1}{2}[$, we find that f is constant on $]0, \frac{1}{2}]$. probability vector $(x, \frac{1}{2}(1-x), \frac{1}{2}(1-x))$ and conclude from the first part of the proof that $f(x) = f(\frac{1}{2}(1-x))$. As $0 < \frac{1}{2}(1-x) < \frac{1}{2}$, we conclude that $f(x) = f(\frac{1}{2}))$. Thus f is constant on $]0, 1[$. By letting $t \to 1$ and assuming suitable smooth behaviour, we conclude that the constant is -1.
We have seen that κ must be the unique solution in $]0, 1[$ to

$$\chi(t)\kappa(t) + t\kappa'(t) = -1$$

for which $\kappa(1) = 0$. $\square$

Tsallis entropy

Below, $\ln_q$ is the *q*-logarithm:

$$\ln_q t = \begin{cases} \ln t \text{ if } q = 1 \\ \frac{1}{1-q}\left(t^{1-q} - 1\right) \text{ otherwise} . \end{cases}$$

Theorem: If π is consistent, (i.e. $\sum z_i = 1$ is ensured), then the world must be one of the Tsallis worlds $\mathcal{W}_q$. If $q < 0$, PMP cannot be realized. If $q \geq 0$ there is a unique generator s.t. PMP holds, viz. the generator $\kappa_q(t) = \ln_q \frac{1}{t}$. The case $q = 0$ is a singular case for which divergence vanishes. When $q > 0$, PMP holds in its full form. The entropy obtained for $q \geq 0$ coincides with Tsallis entropy.

Proof (sketch): Consistency $\Rightarrow$ Tsallis world: easy! Differential equation from previous theorem: $(1-q)\kappa(t) + t\kappa'(t) = -1$ gives $\kappa(t) = \ln_q \frac{1}{t}$. Counterexamples rules out $q < 0$ if PMP is to hold. That PMP does hold when $q \geq 0$ follows as PFI holds (apply geometric/arithmetic mean inequality). $\square$

PART III: Outlook

- information without probability, paradigms of cognition
- belief as a tendency to act
- control derived from belief
- description effort, or description cost
- multiple worlds with the same generator (κ-function)
- the choice of unit (as cost pr. observation)
- "what *can* you know?", feasible preparations
- notions of equilibrium (robustness, Nash equilibrium)
- exponential families, the maximum entropy principle
- information projections, pythagorean theorems