Om sandhed, tro og viden

Flemming Topsøe

Institut for Matematiske Fag
Københavns Universitet
http://www.math.ku.dk/ topsoe med mange manuskripter
se specielt http://www.math.ku.dk/ topsoe/sandhednatfest09.pdf

Naturvidenskabsfestival
21-25 september, 2009

Verden

Det hele er verden , $\mathcal{V}$.
Situationer fra verden vedrører Naturen og dig, Iagttageren.

Naturen bærer (gemmer!) sandheden,
Du søger sandheden, men er henvist til tro .

Med erfaring opnår du viden.
Viden er syntesen af udstrakt erfaring.

situationer

Naturfænomener: solopgang, vejret i morgen, ...
Fra fysikkens verden: tilstanden af en gas i et varmebad, ...
Sundhedssektoren: virker pillen? kommer der en epedemi? ...
Samfund, økonomi: kursernes udsving ...
Den religiøse sfære: forholdet mellem Vorherre og dig, ...
Personlige forhold: elsker hun mig? ...
Spil: er terningen ægte? ...

Naturen og dig

Naturen har ingen bevidsthed, er ikke kreativ - det er du!

Konkretiserede størrelser (instanser) i specifikke situationer:
sandhed: x tro: y viden: z

En lidt anden (?) fortolkning af viden (z):
sådan opfatter du sandheden! (perception)

Fundamental og kritisk antagelse:
viden afledes af sandhed og tro tilsammen
– dvs. der findes en funktion, interaktoren Π, således at

$$z = \Pi(x, y)\,.$$

Interaktoren karakteriserer verden: $\mathcal{V} = \mathcal{V}_\Pi$.

Eksempler på verdener

Den klassiske verden $\mathcal{V}_1$ er karakteriseret ved interaktoren Π_1:

$\Pi_1(x, y) = x$, så her er $z = x$. Her kan sandheden erfares:

det, du ser, er det, der er sandt. (DDSDDS)

Et sort hul $\mathcal{V}_0$ er karakteriseret ved interaktoren Π_0:

$\Pi_0(x, y) = y$, så her er $z = y$, med andre ord:

det, du ser, er det, du tror. (DDSDDT)

Blandinger, f.eks $\mathcal{V}_{\frac{1}{4}}$, $\mathcal{V}_{\frac{1}{2}}$, $\mathcal{V}_{\frac{3}{4}}$, generelt:

Tsallis' verden $\mathcal{V}_q$, karakteriseret ved $\Pi_q(x, y) = qx+(1-q)y$.

noget at tænke over

Opgave 1: Diskuter situationer, hvor de opstillede filosofiske betragtninger virker interessante, måske fordi de peger på forhold, du selv har iagttaget. Du bør forholde dig kritisk til de fremførte tanker. Der er masser af spørgsmål, man kan stille, f.eks. omkring begrebet “virkelighed” - findes “virkeligheden” og, i sammenhæng hermed, er der mening i et begreb som den “absolutte sandhed”?
Hvordan passer begreberne sandhed, tro og viden ind i en religiøs kontekst? Er det det rene vås og misforstået videnskabelighed, hvis man prøver at fortolke begreberne i en anden sammenhæng end en rent videnskabelig?
Og, lidt omvendt, kan forhold vi møder som almindelige mennesker hjælpe os, når vi søger efter modeller til forklaring af rent naturvidenskabelige forhold?

Verdener baseret på sandsynlighed

Verdener, hvor situationer er bestemt ved sandsynligheder over et alfabet. Skematisk:

$\mathbb{A}$	sandhed	tro	viden
.	.	.	.
i	x_i	y_i	z_i
.	.	.	.

(f.eks i $\mathcal{V}_{\frac{1}{4}}$ er $z_i = \frac{1}{4}x_i + \frac{3}{4}y_i$). Et konkret eksempel:

$\mathbb{A}$	sandhed	tro	viden i$\mathcal{V}_1$	viden i$\mathcal{V}_0$	viden i$\mathcal{V}_{\frac{3}{4}}$	viden i$\mathcal{V}_{\frac{1}{4}}$
$\diamondsuit$	$\frac{1}{4}$	$\frac{1}{4}$	0.2500	0.2500	0.2500	0.2500
$\heartsuit$	$\frac{1}{2}$	$\frac{1}{4}$	0.5000	0.2500	0.4375	0.3125
$\clubsuit$	$\frac{1}{8}$	$\frac{1}{4}$	0.1250	0.2500	0.1563	0.2188
$\spadesuit$	$\frac{1}{8}$	$\frac{1}{4}$	0.1250	0.2500	0.1563	0.2188

score funktioner

En score-funktion Φ er en funktion, der i enhver situation angiver dit besvær med at opnå viden (omkostning, pris, nødvendig energi...). Afhænger af x og af y: $\Phi(x, y)$.

Normalt har Iagttager flere score-funktioner at vælge imellem. **Princip:** Vælg Φ, så besværet er mindst, når der foreligger en perfekt match, dvs. $y = x$. Altså: sørg for at $\Phi(x, y) \geq \Phi(x, x)$ og for at "$=$" kun gælder når $y = x$. En sådan funktion hedder en ren score-funktion.

Vi kan opfatte Φ som *prisen på viden* (aktiv fortolkning, aktiv fordi der typisk skal handles, f.eks. udføres forsøg, for at indhente viden) eller som *information, der kan vindes*; mere præcist som et kvantitativt mål for den informationsmængde, der kan vindes (passiv fortolkning).

score-funktioner i stokastiske modeller

$\mathbb{A}$	x	y	z	besvær (pris, energi...)
.	.	.	.	.
i	x_i	y_i	$z_i = \pi(x_i, y_i)$	$\phi(x_i, y_i) = z_i\kappa(y_i) = \pi(x_i, y_i)\kappa(y_i)$
.	.	.	.	.

Det totale besvær er så $\Phi(x, y) = \sum_{i \in \mathbb{A}} \pi(x_i, y_i)\kappa(y_i)$.

Her er $\kappa : [0, 1] \mapsto [0, \infty]$, en funktion, der til en given værdi t (sandsynlighed, du tror, et udfald har), angiver det besvær, $\kappa(t)$, du er villig til, eller nødt til, at acceptere for at finde frem til udfaldet [uddybning gives!].

Vi forlanger om en κ-funktion , at den er "glat" (differentiabel), at $\kappa(1) = 0$ og at $\kappa'(1) = -1$. Det sidste krav er en *normeringsbetingelse* og svarer til et valg af enhed . Vores valg giver naturlige enheder, nats. Havde vi valgt normeringen $\kappa'(1) = \ln \frac{1}{2} \approx -0.6931$, havde vi fået binære enheder (bits).

1 nat$= \frac{1}{\ln 2}$ bit $\approx$ 1.4427 bit.

... fortsat...

Opsummering: To funktioner er centrale: π og κ:

$\pi(s,t)$ angiver den vægt (sandsynlighed) et udfald opfattes at have i din verden, når udfaldet har sand sandsynlighed s og du tror, sandsynligheden er t. Typisk er $\pi = \pi_q$ svarende til en Tsallis verden, men der er andre muligheder (skal kræve, at $\pi(s,s) = s$ for alle s).

Du vælger κ. Når κ er valgt, kan du tilrettelægge observationer således, at ethvert udfald, du tror har sandsynlighed t, medfører et besvær for dig af størrelsen $\kappa(t)$ (OBS: afhænger *ikke* af den sande sandsynlighed!). Mulig fortolkning: κ angiver besværet med at beskrive udfald.

Sætning: For enhver verden $\mathcal{V}_\pi$ er der masser af acceptable score-funktioner, nemlig svarende til forskellige valg af κ, men højst *eet* valg, der giver en ren score funktion!

Beviset er lidt omstændeligt. Jeg springer det over!

Vi gætter en ren score fkt. i klassiske verden $\mathcal{V}_1$

Skal finde κ, så det for alle sandsynlighedsvektorer x og y gælder, at summen $\Phi(x,x)$ er $\leq$ summen $\Phi(x,y)$, se skema:

$\mathbb{A}$	x	y	bidrag til $\Phi(x,x)$	bidrag til $\Phi(x,y)$
1	x_1	y_1	$x_1\kappa(x_1)$	$x_1\kappa(y_1)$
2	x_2	y_2	$x_2\kappa(x_2)$	$x_2\kappa(y_2)$
.	.	.	.	.
i	x_i	y_i	$x_i\kappa(x_i)$	$x_i\kappa(y_i)$
.	.	.	.	.
sum	1	1	$\Phi(x,x)$	$\Phi(x,y)$

Et trick: Vis i stedet, at
summen $\Phi(x,x)+1$ er $\leq$ summen $\Phi(x,y)+1$, se skema:

... fortsat...

$\mathbb{A}$	x	y	bidrag til $\Phi(x,x)+1$	bidrag til $\Phi(x,y)+1$
1	x_1	y_1	$x_1\kappa(x_1)+x_1$	$x_1\kappa(y_1)+y_1$
2	x_2	y_2	$x_2\kappa(x_2)+x_2$	$x_2\kappa(y_2)+y_2$
.	.	.	.	.
i	x_i	y_i	$x_i\kappa(x_i)+x_i$	$x_i\kappa(y_i)+y_i$
.	.	.	.	.
sum	1	1	$\Phi(x,x)+1$	$\Phi(x,y)+1$

Satser på at dette endog gælder "ledvist", dvs. at uligheden

$$s\kappa(s)+s \le s\kappa(t)+t \tag{1}$$

gælder for alle $0 \le s \le 1$ og alle $0 \le t \le 1$. Hold først s fast. Højre-siden i (1) er en funktion af t med mindsteværdi for $t=s$; derfor er der stationært punkt for $t=s$, dvs. $s\kappa'(s)+1=0$. Dette gælder alle s og bestemmer dermed en differentialligning. Løsningen med $\kappa(1)=0$ er funktionen

$$\kappa(t)=\ln\frac{1}{t}\,.$$

... fortsat ...

Vi har gættet en κ-funktion og dermed en score funktion! Men er det en ren score funktion? Vi checker: Er

$$s\kappa(s) + s \leq s\kappa(t) + t \ ?$$

– eller: er

$$s \ln \frac{1}{s} + s \leq s \ln \frac{1}{t} + t \ ?$$

– eller: er

$$s \ln \frac{t}{s} \leq t - s \ ?$$

JA! – det følger af den velkendte ulighed $\ln x \leq x - 1$! Vi konkluderer:

$$\boxed{\text{I } \mathcal{V}_1 \text{ er } \kappa : \ t \curvearrowright \ln \tfrac{1}{t} \text{ den entydigt bestemte rene score funktion!}}$$

Entropi og divergens

Se på abstrakt model med en ren score funktion Φ.
Entropien af x, der betegnes $H(x)$ defineres som den minimale værdi af score funktionen, altså det minimale besvær for dig. Det opnås, når der er en perfekt match, dvs.

$$H(x) = \Phi(x, x)\,.$$

Overskudsværdien sv.t. et givet y kaldes divergensen mellem x og y og betegnes $D(x, y)$, altså

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

Opgave 2: Bestem entropi og divergens i $\mathcal{V}_1$.

score-funktioner i $\mathcal{V}_q$

Opgave 3: Se på Tsallis verden $\mathcal{V}_q$ med et q mellem 0 og 1 ($0 < q < 1$). Der er kun een PMP-score funktion i denne verden. Bestem den og opskriv en formel for den tilhørende entropi, den såkaldte Tsallis entropi .
Vejledning: Gå frem helt som i $\mathcal{V}_1$ og gæt dig frem. Tricket, vi så før, fører til differentialligningen

$$s\kappa'(s) + (1-q)\kappa(s) + 1 = 0\,.$$

Du finder let en konstant-funktion κ_0 som løsning og så ses, at er κ en løsning, opfylder funktionen $f = \kappa - \kappa_0$ en noget simplere differentialligning, som du kan løse. Husk så at pil den løsning ud, der opfylder $\kappa(1) = 0$. Du kan stille dig tilfreds med dette, men bør checke, at uligheden $s\kappa(s) + s \leq s\kappa(t) + t$ virkelig er opfyldt.
Advarsel: Opgaven er ikke let! Vil du vide mere om Tsallis entropi, spørg evt. Google eller Wikipedia.

Hvordan sikrer man sig et ærligt svar?

Tilbage til abstrakt set-up. Ny fortolkning:

"Naturen" kan kommunikere! Så taler vi om en ekspert.
Du spørger eksperten om råd.
Ekspertens viden er x, det afgivne råd er y.

Men hvad nu, hvis eksperten er fristet til at give et råd mod bedre vidende ($y \neq x$)? Med en ren score funktion kan man sikre sig mod dette:

> **Strategi:** Betal eksperten et beløb for overhovedet at få et råd. Indgå en slags forsikring med eksperten, således, at eksperten betaler dig en "straf" af størrelsen $\Phi(x, y)$ såsnart x er kendt.

Det bør afholde eksperten fra at vælge et råd y forskellig fra x.
Opgave 4: Overvej selv dette nærmere, foreslå evt. konkrete situationer, hvor modellen er rimelig.

0'er og 1'er i $\mathcal{V}_1$ (kodning!), binære enheder (bits)

I $\mathcal{V}_1$ giver kodning operationel fortolkning af Φ mv. Eksempler:

$\mathbb{A}$ i	sandh. x_i	tro y_i	din kode	kode lgd.	bidr. t. $\Phi(x, y)$	bidr. t. $H(x)$	bidr. t. $D(x, y)$
♢	0.250	0.250	00	2	0.500	0.500	0.000
♡	0.500	0.250	01	2	1.000	0.500	0.500
♣	0.125	0.250	10	2	0.250	0.375	−0.125
♠	0.125	0.250	11	2	0.250	0.375	−0.125
sum	1	1	-	-	2.000	1.750	0.250

$\mathbb{A}$ i	sandh. x_i	tro y_i	din kode	kode lgd.	bidr. t. $\Phi(x, y)$	bidr. t. $H(x)$	bidr. t. $D(x, y)$
♢	0.250	0.250	00	2	0.500	0.500	0.000
♡	0.250	0.500	1	1	0.250	0.500	-0.250
♣	0.250	0.125	010	3	0.750	0.500	0.250
♠	0.250	0.125	011	3	0.750	0.500	0.250
sum	1	1	-	-	2.250	2.000	0.250

... fortsat

Opgave 5: Prøv selv et forklare nærmere! Det bør føre dig ind på spændende ting omkring kodning.
Du kan let finde ting herom på nettet. Desuden vil din lærer kunne låne/give dig et eksemplar af en lille bog om informationsteori, der indeholder meget, meget mere om koder etc., men altsammen i den klassiske verden $\mathcal{V}_1$.
Bemærkning: Det er uvist om der findes en lignende operationel kodningsteoretisk fortolkning af entropi mv. i andre verdener, specielt i en af Tsallis' verdener, $\mathcal{V}_q$. Hvis du finder en sådan fortolkning, vil mange forskere være ekstremt interesserede i at høre nærmere!

Tak for nu!