NORTH ATLANTIC TREATY ORGANIZATION

SCIENCE AND TECHNOLOGY ORGANIZATION

AC/323()

www.sto.nato.int

STO TECHNICAL REPORT

PUB REF STO-MP-SAS-114-PPH

ANNEX H
Information Gain and Approaching True Belief

Jonas Clausen Mork

Information gain and approaching true belief (and something about uncertainty and scoring rules as well)

Jonas Clausen Mork, FOI (SWE)

December 6, Copenhagen

SAS-114, fourth meeting

Another perspective on information

- The ”location” of information
- The (standard) ”infon” view: physical objects (or configurations of such objects) have information values (contain information), typically stuff outside minds, like books, radio signals or symbols on a screen
- Example: this note contains 2 bits of information
- Another view: nothing outside of minds has information values except in a derived sense – what carries information are belief state transitions
- Example: I just made a 4-bit belief change

A derived sense of information contained

- Roughly: a given object has information content w.r.t. a specified belief state (possibly part of a mind) to the extent that the object (typically when perceived) would (counterfactually) induce a change in that state
- Example: this book contains 2 bits of information w.r.t. Bob's current belief state, since it would induce a 2-bit information gain, were he to read it (but he hasn't yet…)
- Also, if Bob read it and updated his beliefs as much as the book has potential to do so, the book no longer contains information. For Bob.
- Information content changes over time
- Information content differs between observers

FOI

But, what is information again?

- Let's call the amount of information I, and then consider some probability function *p*
- Different views on what *p* is in different analyses
- Improbable things (events) ”contain” a lot: I is big if *p*(e) is small
- It has something to do with truth: I is small for false statements and big for true statements
- People can gain it and lose it: ΔI can be positive or negative
- Experiments can provide it, and at worst give us none (really?): ΔI(experiment)$\geq$ 0

A Bayesian representation of belief state

- Subjective probability distributions over sets of hypotheses
- Discrete case
- Hypothesis set: $\mathbf{H} = \{h_1, h_2, \ldots, h_n\}$, at most one is true
- Probability distribution p over **H**
- Belief change: $p \rightarrow q$, p is prior, q is posterior
- Transition not necessarily rational
- So, the p right now is a subjective probability

Two analyses of information gain, ex post

- Information gain of a (after a) transition
- Objective information gain: belief (doxastic) movement towards true belief
- Subjective information gain: expected (with posterior *q*) objective information gain
- Objective information gain (bits):
  - $I(p,q) = log_2(q(t)/p(t))$, where *t* is the true hypothesis
- Subjective information gain (bits):
  - $I^*(p,q) = \sum q(h) log_2(q(h)/p(h))$
- The amount of objective information gained is typically not possible to assess for a believer, whereas the subjective information gain typically is

FOI

Objective information gain, example

- Three men are suspected of stealing: Adam, Bob and Caesar
- Bob did it
- The investigator, Joan, has the following belief state:
  - p(Adam) = 0.22, p(Bob) = 0.17, p(Caesar) = 0.61
- Joan conducts interrogations with the suspects, after which her belief state is the following:
  - q(Adam) = 0.28, q(Bob) = 0.27, q(Caesar) = 0.45
- Joan (objectively) gained $\log_2(0.27/0.17) \approx 0.67$ bits

FOI

Subjective information gain, example

- Same setting, same transition
  - $p(\text{Adam}) = 0.22$, $p(\text{Bob}) = 0.17$, $p(\text{Caesar}) = 0.61$
  - $q(\text{Adam}) = 0.28$, $q(\text{Bob}) = 0.27$, $q(\text{Caesar}) = 0.45$
- Joan (subjectively) gained $0.28 \log_2(0.28/0.22)+0.27 \log_2(0.27/0.17)+0.45 \log_2(0.45/0.61) \approx 0.08$ bits
- So, she would perhaps consider the interrogations relatively uninformative, whereas, in fact, they did bring her (beliefs) closer to true belief

Some consequences of this view

- Dataless information gain (e.g. in mathematics or reasoning in intelligence analysis)
- Irrational or unjustified information gain
- Unconvincing data are just data, not information (in the derived sense)

FOI

Something about uncertainty

- Subjective probability is not a measure of uncertainty, it is a measure of belief
- Uncertainty is measured in the same unit as information, e.g. bits (base 2), nats (base *e*) or Hartleys (base 10)
- Uncertainty is the expected information gain upon reaching certainty (not necessarily about the truth…)
- For Bayesian representations, the entropy of *p* is the only reasonable measure of uncertainty

Proper scoring rules, information gain and uncertainty

- Brier scores are popular in forecasting, strictly proper and with a useful decomposition
- Benedetti (2010) has shown how the Brier score depends on probabilities assigned to non-occurring events (*nonlocality*), which is undesirable
- Two forecasters who assigned the same probability to the event that occurred can get different scores
- *Additivity*, *locality, differentiability* and *strict propriety* uniquely picks out the logarithmic scoring rule

Proper scoring rules, information gain and uncertainty, contd.

- The logarithmic scoring rule is the negative (subjective) information gain upon certainty, i.e. you're penalized if observing (becoming certain of) the event provided you with a lot of information
- The score coincides with the negative objective information gain if what you become certain of is also the truth – a subtle difference, but not unimportant for a scoring regime (issues will close on certainty, not truth as such)

FOI

Proper scoring rules, information gain and uncertainty, contd.

- The expected score for a given problem is, then, the negative uncertainty (p-expectation of negative subjective information gain upon certainty)
- The expected score change after a *soft update* (not to certainty) is (almost) the negative subjective information gain (q-expectation of negative objective information gain)
- For this to work out, there needs to be an overlap between the probability that a given hypothesis is *true* and the probability that a given hypothesis is the one that a scoring regime will accept certainty of at some point
- If not, it of course gets more complicated…

FOI

Papers

- Clausen Mork (2015) ”Information gain and approaching true belief”, *Erkenntnis* 80(1):77-96
- Clausen Mork (2013) ”Uncertainty, credal sets and second order probability”, *Synthese* 190(3):353-378
- Benedetti (2010) ”Scoring Rules for Forecast Verification”, *Monthly Weather Review* January 2010:203-211

FOI