Pragmatic Information for Cognitive Agents

05.11.2008

Workshop on Coordination of Agents

King’s College London

Peter beim Graben
Dept. of Clinical Language Sciences
University of Reading
p.r.beimgraben@reading.ac.uk
Information Theory

“Relative to the broad subject of communication, there seem to be problems at three levels. Thus it seems reasonable to ask, serially:

**Level A.** How accurately can the symbols of communication be transmitted? (The technical problem.)

**Level B.** How precisely do the transmitted symbols convey the desired meaning? (The semantic problem.)

**Level C.** How effectively does the received meaning affect conduct in the desired way? (The effectiveness problem.)”

W. Weaver in Shannon & Weaver (1949)
von Weizsäcker’s Proposals

1. Pragmatic information measures effect upon the receiver
2. Pragmatic information is minimal for totally novel and for completely confirmed messages
3. Pragmatic information requires a non-classical (quantum like) description

E. v. Weizsäcker & C. v. Weizsäcker (1972)
Effectiveness

“Information [...] should act. By definition they act upon their receivers and change them informationally. In particular, after the arrival of a message the receiver's expectation probability of a related message will usually not be the same as before.”

“If information acts successfully, it changes the basis of its own quantification.”

E. v. Weizsäcker & C. v. Weizsäcker (1972)
Novelty and Confirmation

**novelty:**
Minä vien alituiseen vaimoani joka paikkaan, mutta aina hän löytää tiensä sieltä takaisin.

**confirmation:**
when I was young, I wanted to grow up to be somebody. Now that I am older – I wish I had been more specific.

E. v. Weizsäcker & C. v. Weizsäcker (1972)
Non-Classicality

quantum theory:
complementary observables are Fourier pairs

E. v. Weizsäcker & C. v. Weizsäcker (1972)
Effectiveness

“Dynamic semantics is called ‘dynamic’ because it assumes that the meaning of a sentence is not its truth condition but rather its impact on the hearer.”

Gärdenfors’ formalism of belief models for cognitive agents

Kracht (2002)
Dynamic Semantics

- cognitive agent
- epistemic states (belief states) \( x, y, z \in X \)
- state transitions (semantic operators)
  \[ A, B, C : \begin{cases} 
  x & \rightarrow & y \\
  X & \rightarrow & X 
  \end{cases} \]
- composition
  \[ (AB)(x) = (A \circ B)(x) = A(B(x)) \]
- identity
  \[ \top(x) = x \]

Gärdenfors (1988)
Propositions I

- conjunction
  \[ A \land B = AB = BA \]
- idempotence
  \[ A^2 = AA = A \]

\[ A, B, C \in \mathcal{P} \subset \text{Mor}(X) \]

Gärdenfors (1988)
Propositions II

- acceptance

A proposition \( A \in \mathcal{P} \) is accepted in state \( x \in X \) if

\[
A(x) = x
\]

- logical consequence

Let \( A, B \in \mathcal{P} \). \( B \) is called logical consequence of \( A \), if \( A \land B = B \land A = A \).

interpretation: If state \( y \) accepts proposition \( A \), then \( y \) accepts \( B \), too.
Belief Models

$(X, \mathcal{P})$ with those properties is a belief model with state space $X$ and the set of propositions $\mathcal{P}$. By further axioms, all logical connectives can be introduced, such that $(X, \mathcal{P})$ is interpretable as classical propositional logic. The meaning of a proposition $A \in \mathcal{P}$ is its impact upon an agent’s epistemic state space $X$.

Hence: meaning is context-dependent: 1. w.r.t the cognitive agent, 2. w.r.t. her current epistemic state (beliefs, knowledge, …)

Gärdenfors (1988)
Bayesian Belief Models

statistical (Bayesian states):
probability distributions over propositions
\( p : \mathcal{P} \rightarrow [0, 1] \subset \mathbb{R} \); \( p(A) \) probability for accepting \( A \).

limiting cases:
\( p(A) = 0 \): \( A \) not accepted in state \( p \).
\( p(A) = 1 \): \( A \) certainly accepted in state \( p \).
Conditionalization

Let \( p : \mathcal{P} \rightarrow [0, 1] \subset \mathbb{R} \) be a statistical state and \( A \in \mathcal{P} \) a proposition. \( A \) acts as a semantic operator on \( p \) by means of the map

\[
A : p \mapsto p_A
\]

\[
p_A(P) = \frac{p(PA)}{p(A)} = p(P|A)
\]

for all \( P \in \mathcal{P} \), if \( p(A) > 0 \).

henceforth, \( A \) is accepted: \( p_A(A) = 1 \).
Pragmatic Information

A measure for pragmatic information is given by the mean information gain supplied by the transition from the prior state \( p \) to the conditionalized state \( p_A \), after accepting the proposition \( A \).

\[
K(p_A, p) = \sum_{P \in \mathcal{P}} p_A(P) \log \frac{p_A(P)}{p(P)}
\]

\[
S_p(A) = \begin{cases} 
K(p_A, p) & : \ p(A) > 0 \\
0 & : \ p(A) = 0 
\end{cases}
\]
Novelty and Confirmation

Call a proposition $A \in \mathcal{P}$ (totally) novel, if no state $p$ can be conditionalized by $A$.

$p_A = p$

$S_p(A) = 0$

Call a proposition $A \in \mathcal{P}$ (completely) confirmed, if $A$ is logical consequence of an already accepted proposition $B$.

$S_{p_B}(A) = 0$
Texts

Example: Let $A, B \in \mathcal{P}$ be propositions. Call $T = ABBA = A \land B \land B \land A \in \mathcal{P}$ a text.

propositions commute

$T = ABBA = AABBB = BBAA$

relative novelty not definable

propositions are idempotent

$T = A^2B^2 = AB = B^2A^2 = BA$

relative confirmation not definable
Non-Classicality

Consider semantic operators \( U, V \in \text{Mor}(X) \setminus \mathcal{P} \).

Generally

\[ UV \neq VU \quad \text{“complementarity”} \]

\[ U^2 \neq U \quad \text{no idempotence} \]

generalized conditionalization

\[ p_U(V) = \frac{p(VU)}{p(U)} = p(V | U) \]
Statistics

- variance of a semantic operator $U$ in state $p$:
  \[ \text{vax}_p(U) = p(U^2) - p(U)^2 \]
- covariance of a semantic operator $V$ in state $p$ given $U$:
  \[ \text{cox}_p(V|U) = p(VU) - p(V)p(U) \]

properties:

- for propositions $A \in \mathcal{P}$ holds $\text{vax}_p(A) = 0$
- if $\text{vax}_p(U) > 0$, then $p_U(U) > p(U)$ (conviction)
- if $\text{vax}_p(U) < 0$, then $p_U(U) < p(U)$ (propaganda)
Novelty and Confirmation

Call an operator $V \in \text{Mor}(X)$ (relatively) novel in state $p$ w.r.t. an operator $U \in \text{Mor}(X)$, if
\[
\text{cox}_p(V|U) = 0
\]

Call an operator $V \in \text{Mor}(X)$ (relatively) confirmed by the operator $U \in \text{Mor}(X)$ in state $p$, if
\[
\text{cox}_p(V|U) > 0
\]

(For $\text{cox}_p(V|U) < 0$ we have relative disconfirmation)
Pragmatic Information

The generalized measure of pragmatic information is the averaged information gain attributed to the transition from state $p$ to state $p_U$, after accepting operator $U$.

\[
K(p_U, p) = \sum_{V \in A \subseteq \text{Mor}(X)} p_U(V) \log \frac{p_U(V)}{p(V)}
\]

\[
S_p(U) = \begin{cases} 
K(p_U, p) & : \quad p(U) > 0 \\
0 & : \quad p(U) = 0
\end{cases}
\]
Decomposing Pragmatic Information

Decomposing the sum into three parts:

\[ S_p(U) = \sum_{V: \text{cox}(U|V) < 0} + \sum_{V: \text{cox}(U|V) = 0} + \sum_{V: \text{cox}(U|V) > 0} \]

- disconfirmation
- novelty
- confirmation

= 0

Compare: Kornwachs and von Lucadou (1982): \[ S = E \times B \]
Summary

1. Pragmatic information measures the impact of semantic operators upon epistemic states of a cognitive agent.

2. In the limiting cases of total novelty / complete confirmation, pragmatic information vanishes. It consists of confirmation and disconfirmation.

3. Relative novelty / confirmation require non-classical operators such as in quantum theory.

E. v. Weizsäcker & C. v. Weizsäcker (1972)
Applications

1. semantic information
2. relevance
3. neural correlates
Semantic Information

Boolean algebra \( \mathcal{P} = \{\top, \bot\} \)

apriori distribution \( p(\top) = \beta; \quad p(\bot) = 1 - \beta \)

aposteriori distribution \( p_{\top}(\top) = 1; \quad p_{\top}(\bot) = 0 \)

pragmatic information
\[
S_p(\top) = -\log \beta = \inf(\top)
\]

Bar-Hillel & Carnap (1953, 1964)
Relevance

Average pragmatic information of all operators

\[ S_p = \sum_U p(U) S_p(U) : \]

\[ S_p = \sum_{U,V} p(UV) \frac{p(UV)}{p(U)p(V)} \]

= mutual information = relevance = utility function in evolutionary game theory of interacting agents

Polani et al. (2001)
Weinberger (2002)
von Rooij (2004)
Neural Correlates

P300 amplitude correlates with
- subjective a posteriori probability
- task-relevance
- equivocation
- mutual information

Ruchkin & Sutton (1978)
Sutton (1979)
Johnston (1979)
Donchin (1981)
Sutton & Ruchkin (1984)
Thank you for your attention!