

DECISION THEORY OF A LANGUAGE UNIVERSAL

Arthur Merin

IMS, University of Stuttgart, Azenbergstr. 12, 70174 Stuttgart, Germany
 Abstracts of the Tenth International Congress of Logic, Methodology and Philosophy of
 Science; Florence, August 19-25 1995, [M.L. Dalla Chiara et al. eds.]

Section 7: Probability, Induction and Decision Theory

Customs of conduct and cognition, passed on whichever way, are what human languages instantiate or reflect. A candidate for a universal of grammar, concerning the conjunction ‘but’ and its equivalents, shows that constraints on induction condition language structure in unexpected ways.

Locke and Frege noted that ‘ A but B ’ intimates B unexpected given A . Write: B is A -unexpected. By contrast, R. Fogelin in 1967 and O. Ducrot in 1973 found argumentation: A supports an H , but is overridden by B supporting $\neg H$. Write: A, B are H -contrary, with B overriding. *Example*: ‘Kim is a doctor but she lives in R.’ where R. is notorious for low GNP and $H =$ ‘Kim is rich’.

Explicate now in terms of stochastic relevance $r_H(E)$ of a proposition E to another, H . Using the log-likelihood-ratio ($r_H(E) =_{df} \log[p(E|H)/p(E|\neg H)]$), $r_H(E) = 0$ when E is irrelevant to H ; and if E and F are p-independent conditional on H and $\neg H$, $[(E \perp F|H, \neg H)]$ then $r_H(EF) = r_H(E) + r_H(F)$. (For our application treat p as representing Addressee’s and/or third parties’ epistemic state, as ostensibly imputed by Speaker.)

Proposition 1: If $(A \perp B|H, \neg H)$ then B is A -unexpected [$p(B|A) < p(B)$] iff A and B are H -contrary [$\text{sgn}(r_H(A)) = -\text{sgn}(r_H(B)) \neq 0$]. (Cp. Reichenbach)

Proposition 2: If $r_H(A) > 0$ and $H = \neg B$ and $AB \neq \text{FALSE}$ and r given by a strictly coherent probability function p then (i) $r_H(B) < 0$; (ii) $p(B|A) < p(B)$; (iii) $r_H(AB) = r_H(A) + r_H(B)$; (iv) $-r_H(B), -r_H(AB) > r_H(A)$.

Remark: The special case $H = \neg B$ of H -contrariness (making B A -unexpected) together with compatibility of A, B is a prime candidate for an interpretive default. It yields all basic empirical properties of ‘but’ by purely qualitative reasoning and avoids introducing new atoms into the propositional universe of discourse. *Example*: ‘Kim is poor but honest’.

Sharpening a syntactic observation of Zellig Harris, note next that ‘Kim walks but {Kim/(s)he_K} talks’ is acceptable in suitable contexts. However, ‘Kim walks but Sandy walks’ never is. Yet ‘Kim walks but Sandy {also walks/walks too}’ is again always acceptable. On present evidence, any language having equivalents of ‘but’ and ‘also’ obeys the pattern.

How to explain the apparent asymmetry of subject (particular) and predicate (universal)? Let p be defined on a monadic first order quasi-English fragment.

Definitional Fact: If p satisfies non-negative instantial relevance [NNIR] (Carnap) then Qb cannot be Qa -unexpected. By contrast, NNIR does not rule out Pc being Qc -unexpected nor, say, Qd -unexpected. *Remark*: Yet A -unexpectedness of B is not generally necessary for acceptability of ‘ A but B ’.

Hypothesis: Every natural language L obeys a Default Satisfiability Principle [DSP]: Let L have a default condition Γ on preferred interpretation of a class $\Sigma \subseteq L$ of expressions. Let $\sigma \in \Sigma$. If σ satisfies Γ for no context of non-degenerately empirical use (e.g. $H: 0 \neq r_H(\sigma) \neq \pm\infty$), then σ is unacceptable [write: $^*\sigma$]. *Example:* If $(X \perp Y|H, \neg H)$ is a default presumption [call it CIP], so additivity of r is, then $^*‘A \text{ and } A’$ for any aspectually stative sentence A .

Corollary: Given DSP, CIP (CIP is also instantiated by the default favoured by Prop. 2), if the probability underlying r obeys NNIR, then $^*‘Qa \text{ but } Qb’$.

Can NNIR, for the linguistically pertinent p -function, be motivated as other than a brute, albeit behaviourally or evolutionarily plausible fact? Yes. Argumentation involving H -contraries uses Universal Probabilistic Instantiation [UIP]. *Example:* ‘Kim is a doctor but she lives in a deprived area’. Argument: $\forall x[\text{doctor}(x) \rightarrow p(\text{rich}(x)) = \alpha \ (1 > \alpha \gg 0.5)]$, $\text{doctor}(\text{Kim}) \models p(\text{rich}(\text{Kim})) = \alpha$; and similarly for the second conjunct.

Fact (Skyrms): Validity of UIP requires that no p -relevant information be conveyed by the individual constant. Names must be treated as properly arbitrary and thus p must be symmetric.

Hypothesis: Natural language semantics satisfies an Embeddability Desideratum [ED]. Models or Discourse Representation Structures (Kamp) generated by a small universe of individuals or their discourse referents should be embeddable in larger models of arbitrary finite cardinality.

Theorem (Savage-Kemeny-Gaifman-Humburg): A symmetric probability function p on finite models m extends to symmetric p' on extensions m' of arbitrary finite size only if p satisfies NNIR.

Corollary: ED and validity of UIP imply NNIR.

What explains the healing power of ‘also’? That it transform p suitably; if by conditioning, then coherently. Note now: in ‘Kim fell and she (also) broke her arm’, ‘also’ removes intimations of causality. And traditionally, ‘ Qa and also Qb ’ marks Qa as epistemically presupposed. But a probability p is an epistemic state. *Definition* (crude): A is presupposed in p iff $\forall X [p(X) = p(X|A)]$.

Corollary: If A is presupposed in p , then $p(B|A) = p(B)$ (independence).

Partial Definition: A is causal for B only if $p(B|A) > p(B)$ (cp. Reichenbach).

Corollary: If A is presupposed, A lacks causal responsibility for B .

Proposition 3: If (i) $p' =_{df} p(\cdot|A)$ and (ii) $0 < p(AB\neg H)$ and r induced by p' , then $r_H(AB) = r_H(A) + r_H(B)$ (r -additive). *Data:* Note a similar decausalizing effect of ‘plus’ or stressed ‘and’, near-synonyms to ‘and also’.

Hypothesis: ‘also’ overrides the A -unexpectedness default for ‘ A but B ’ by conditioning on A . *Evidence:* ‘ A but also B ’ intimates or favours an $H \neq \neg B$.

Outlook: If language qua Humboldt’s *energeia* is a mode of action, one should expect decision theory to inform aspects of its basic structure. The data on ‘but’ might even cast some new light on Ramsey’s claim that the distinction between universals and particulars is purely a matter of language.