

# Probabilistic Semantics of Ethical Reasoning

by Sven Nilsen, 2018

*In this paper I represent an equation that relates probability bias and confidence factor of language context-implication with ethical judgements, based on the theory of ethics as rational reasoning with granular judgements about the world. Together with general path semantical reasoning, this gives a foundation of probabilistic semantics of ethical reasoning, expressed in abstract path semantics.*

Given a two statements `A` and `B` in path semantics, the following equation predicts ethical judgements `E(A)`, `E(A ∧ B)` and `E(A ∧ ¬B)` related by probabilistic semantics:

$$E(A) = C(A \rightarrow B) \cdot E(A \wedge B) + C(\neg(A \rightarrow B)) \cdot (P(B|A) \cdot E(A \wedge B) + P(\neg B|A) \cdot E(A \wedge \neg B))$$

The equation predicts concretely, but is about ethical judgements that are overloaded, so its grounding of truth is within the domain of abstract path semantics, a sub-field of esoteric path semantics.

$E(A)$	An ethical judgement of the statement `A`
$E(A \wedge B)$	An ethical judgement of the combined statements `A ∧ B`
$E(A \wedge \neg B)$	An ethical judgement of the combined statements `A ∧ ¬B`
$C(A \rightarrow B)$	The confidence factor that `A` implies `B`
$C(\neg(A \rightarrow B))$	The confidence factor that it is unknown whether `A` implies `B`
$P(B A)$	The probability bias of `B` given `A`
$P(\neg B A)$	The probability bias of `¬B` given `A`

The inversion for the confidence factor (overloaded since probabilistic implication is undefined):

$$C(\neg(A \rightarrow B)) = 1 - C(A \rightarrow B)$$

The inversion of for probability is standard notation:

$$P(\neg B|A) = 1 - P(B|A)$$

The equation is invariant under the two substitutions:

$$\begin{aligned} C(A \rightarrow B) \cdot E(A \wedge B) &\Leftrightarrow C(A \rightarrow \neg B) \cdot E(A \wedge \neg B) \\ C(\neg(A \rightarrow B)) &\Leftrightarrow C(\neg(A \rightarrow \neg B)) \end{aligned}$$

The substitution is selected by which confidence factor falls within normalized range.

Alternative forms, used to derive confidence factor or probability bias:

$$C(\neg(A \rightarrow B)) = (E(A) - E(A \wedge B)) / ((P(B|A) \cdot E(A \wedge B) + P(\neg B|A) \cdot E(A \wedge \neg B)) - E(A \wedge B))$$

$$P(B|A) = ((E(A) - C(A \rightarrow B) \cdot E(A \wedge B)) / C(\neg(A \rightarrow B)) - E(A \wedge \neg B)) / (E(A \wedge B) - E(A \wedge \neg B))$$