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:

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

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`
$E(A \land B)$	An ethical judgement of the combined statements `A ∧ B`
$E(A \land \neg B)$	An ethical judgement of the combined statements ` A ∧ ¬B`
$C(A = B)$	The confidence factor that `A` equals `B`
$\neg C(A = B)$	The confidence factor that it is unknown whether `A` equals `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:

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

The inversion for probability is standard notation:

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

The equation is invariant under the two substitutions (red replaced with blue and vice versa):

$$C(A = B) \cdot E(A \land B) \Leftrightarrow C(A = \neg B) \cdot E(A \land \neg B)$$
$$\neg C(A = B) \Leftrightarrow \neg C(A = \neg B)$$

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

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

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

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