

Natural Likely

by Sven Nilsen, 2019

The use of words “likely” and “unlikely” in natural languages can be formalized as following:

likely \Rightarrow P `P` is likely

Q \Rightarrow \neg likely `Q` is unlikely

Notice that the inversion rule for likely and unlikely is different from that of `true` and `false`. This rule is because of implication.

$(\text{likely} \Rightarrow P) = (\neg P \Rightarrow \neg \text{likely})$

For example, assume it is likely that a smart AI agent survives:

likely \Rightarrow (smart \Rightarrow survive)

If the AI is smart, it will likely survive:

smart \Rightarrow (likely \Rightarrow survive)

If the AI did not survive, it was likely not smart:

\neg survive \Rightarrow (likely \Rightarrow \neg smart)
 \neg survive \Rightarrow (smart \Rightarrow \neg likely)

However, one can not prove the following:

\neg smart \Rightarrow (likely \Rightarrow \neg survive)

There can be other ways that makes the AI agent likely to survive besides being smart.

The natural use of words “likely” and “unlikely” is not isomorphic to Logic.
One can not give Logic a different interpretation by swapping `true/false` with `likely/unlikely`.