**Propositional Logic**

Q: Can LMs learn the meaning of PL with form-only pretraining? ✓

\[
\begin{align*}
(F\land T) &= (T\land T) \\
(T\land F) &\land F = (F\land (T\land T)) \\
(T\land F) &\land (F\land F) = (T\land T)
\end{align*}
\]

Complete

\[
\begin{align*}
(F\land F) &\lor (T\land F) &= \text{[MASK]}
\end{align*}
\]

Probe

\[
\begin{align*}
(F\land F) &\lor (T\land F) &= \text{[MASK]}
\end{align*}
\]

Hypothesis: Transparency, i.e., context-independence

Transparent language: Learner that can learn its meaning from “assertions”

**Propositional Logic’**

Q: Does transparency affect meaning learnability? ✓

Complete

\[
\begin{align*}
(F\land F) &\lor (T\land F) &= \text{[MASK]}
\end{align*}
\]

Probe

\[
\begin{align*}
(F\land F) &\lor (T\land F) &= \text{[MASK]}
\end{align*}
\]

**NL: Referential Opacity**

Q: Are natural language LMs sensitive to referential opacity? ❌

Lois Lane believes Superman is a hero.

Lois Lane believes Clark Kent is a hero.

纸

Bootstrapping

- symmetry: 50.5
- reflexivity: 92.7
- re/flexivity: 50.3

Significance Test

p-values

GPT-2: 0.66
BERT: 0.57

Permutation

GPT-2: 0.64
BERT: 0.45

LMs can learn the meaning of some languages with form-only pretraining, but doesn’t do well with non-transparent languages.