
Multilingual Physical Reasoning Dataset Shared Task: Kannada

Emile Anand ¹

1. Constructing the Dataset

We provide a dataset of physical reasoning in the Kannada language. In the common paradigm of using paired labels for post-training language models, we provide a carefully hand-curated manually-annotated physical commonsense reasoning evaluation dataset of 100 examples in Kannada. We have supplemented each example with english translations, and our examples reflect some cultural aspects of Karnataka, the Indian state that speaks this language, as well as physical and logical aspects of every-day objects.

For instance, some of the cultural questions asked are (in the English translation), “WHAT IS THE TASTE OF NEEM LEAVES? A) BITTER, B) SWEET”, or “WHAT HAPPENS IF AN AKKI ROTTI IS LEFT FOR A LONG PERIOD OF TIME ON A HOT TAVA? A) IT HARDENS, B) IT SOFTENS”.

In contrast, some of the physical questions asked are “WHEN WALKING IN SOME ROADS IN JAPAN, ELECTRICITY IS GENERATED BY WALKING ON THE ROAD. THIS IS DUE TO A TRANSFER OF ENERGY, WHERE THE FINAL KIND OF ENERGY GENERATED IS ELECTRICAL. WHAT IS THE INITIAL KIND OF ENERGY? A) KINETIC, B) CHEMICAL” or “WHY DO RAINDROPS SLIDE DOWN QUICKLY ON SLOPED TILE ROOFS INSTEAD OF SOAKING IT IN? A) GRAVITY PULLS WATER DOWN SMOOTH INCLINED SURFACES, B) TILES REPEL WATER DUE TO MAGNETIC FORCE”.

The style of questions follows the above examples for the most part, and abides by the standard methodology of communicating in the Kannada language.

¹School of Computer Science, GeorgiaTech, Atlanta, GA, USA.
Correspondence to: Emile Anand <emile@gatech.edu>.