MRL 2025 shared task: physical commonsense reasoning datasets - The Faroese Submission

Authors: Annika Simonsen (University of Iceland), Iben Nyholm Debess (University of the Faroe Islands), Uni Johannesen (University of the Faroe Islands)

All entries in the dataset have been written manually by the authors: three native speakers of Faroese with a master's degree in language or linguistics. Two of the authors are also PhD students in NLP. We did not use LLMs or any other similar tool to write our examples. All examples came from our heads.

Each entry was initially written by one author and subsequently validated by the other two authors. This validation process ensured that examples focused appropriately on physical reasoning while maintaining appropriate difficulty levels. We wanted the examples to be challenging enough but not so difficult that we ourselves could not solve them. During validation, we also addressed minor overlaps that occurred since authors worked independently during the initial creation phase. If someone's sample had been commented as needing to be revised, the author who wrote the sample would then correct their sample and another author validated again.

We followed the task guidelines by making the two candidate solutions as similar as possible, typically differing by only one or two words. A small number of samples have greater dissimilarity between solutions in order to keep the samples interesting and diverse.

The dataset comprises 100 entries in TSV format with the following columns: prompt, solution0, solution1, label, fo-specific, author, validated, en-prompt, en-solution0, and en-solution1.

37 of the entries are specific to Faroese culture, or environment, while 63 entries have topics that are more general or universal regarding physical reasoning. Admittedly, many of them are about food (28 out of 100). The Faroese-specific examples are culturally relevant because they focus on elements central to Faroese life: food preparation and preservation techniques, weather patterns and their physical effects, traditional Faroese clothing, wool and knitting, and the geography of the Faroe Islands. These Faroese physical reasoning scenarios that may not translate directly to other cultures, so we believe they are interesting to use for this task.

The prompts and solutions were translated into English using the neural machine translation (NMT) model *barbaroo/nllb_200_600M_fo_en* (Scalvini, 2024), then manually corrected by one author. The NMT model made several errors due to the large amount of Faroese-specific terms in the text. The English translations are not direct translations intended for use as data, but rather descriptions to help non-Faroese speakers understand the content of the Faroese dataset.

We acknowledge that defining the exact line between physical commonsense reasoning and general commonsense reasoning was very difficult. We have made our best effort to

maintain focus on physical properties and interactions, but recognize that some examples may exist in a gray area where this distinction becomes blurred.

References:

Scalvini, B. (2024, May 16). nllb_200_600M_fo_en [Neural machine translation model]. Hugging Face. https://huggingface.co/barbaroo/nllb_200_600M_fo_en