Commonsense Reasoning with Implicit Knowledge in Natural Language

*Pratyay Banerjee, *Swaroop Mishra, *Kuntal Pal, *Arindam Mitra, Chitta Baral

Can we reason about Abductive, Social and Physical Interaction situations ?

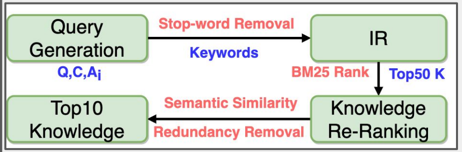
Full Version : https://openreview.net/pdf?id=a4-fFL7aCi0

Abductive NLI	Social IQA	Physical IQA
Obs1: Jim was working on a project. ✓ Jim found he was missing an item. ✗ Jim needed a certain animal for it. **Obs2**: Luckily, he found it on a nearby shelf **Knowledge**: Peyton eventually found it before Peyton needed to determine that something is missing. Kendall never found it, as a result Kendall wants to lodge a missing complaint.	**Context**: Remy was an expert fisherman and was on the water with Kai. Remy baited Kai's hook. **Question**: What will Remy want to do next? ✓ cast the line ✗ put the boat in the water ✗ invite Kai out on the boat **Knowledge**: Alex baits Pat's hook as a result others want to cast their line.	**Goal**: When doing sit-ups: ✓ place your tongue in the roof of your mouth. It will stop you from straining your neck. ✗ place your elbow in the roof of of your mouth. It will stop you from straining your neck. **Knowledge**: How to Do Superbrain Yoga. Place your tongue on the roof of your mouth.

- Can we answer commonsense questions ?
- Can we improve our answers with implicit knowledge in natural language form ?
- How can we find relevant knowledge ?
- How to reason with implicit relevant knowledge ?
- What kind of training strategies can we use ?
- How do they compare with respect to large/complex models ?

Knowledge Retrieval

aNLI <= ROC Stories + Story Cloze
SIQA <= Atomic
PIQA <= Wikihow



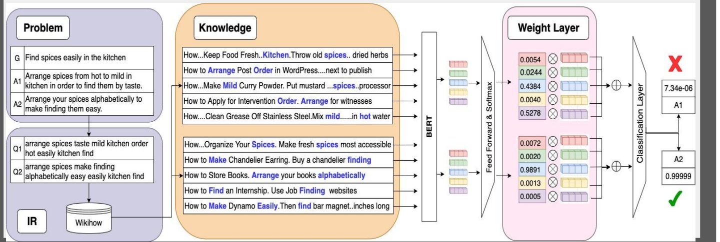
Knowledge Infusion Models

CONCAT [CLS] K_i [SEP] Q A_i [SEP]
PARALLEL-MAX [CLS] K_{ij} [SEP] Q A_i [SEP] → max (CLS for each A_i)
SIMPLE-SUM [CLS] K_{ij} [SEP] Q A_i [SEP] → sum (CLS for each A_i)
WEIGHTED-SUM [CLS] K_{ij} [SEP] Q A_i [SEP] → wtd-sum (CLS for each A_i)

Training Strategies

REVISION T(Bert + RoBERTa) PT (MLM) on K_D => FT on D
OPENBOOK T(Bert + RoBERTa) FT on D+S, S is a subset of K_D
REVISION+OPENBOOK T(Bert + RoBERTa) PT (MLM) on K_D =>FT on D+S

Experiment : Flow Diagram



Results and Analysis

- Revision with Openbook best across 3 datasets for each of 4 knowledge fusion modes
- WS with BERT and RoBERTa is better than baseline BERT and RoBERTa
- Better than huge parameterized model (T5 in aNLI)
- WS shows better performance than complex Graph based models

Types of Error	aNLI	SIQA	PIQA
Annotation	41%	38%	10%
Model Prediction	48%	27%	29%
Distracting Knowledge	11%	35%	61%

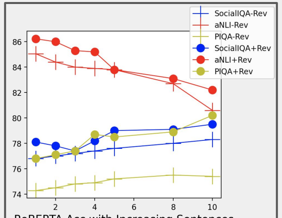
Error %

Knowledge	aNLI	SIQA	PIQA
Explicitly Present	14%	11%	10%
Implicitly Present	55%	59%	51%
Fully Irrelevant	31%	30%	39%

Correct predictions %

		BERT				RoBERTa			
Dataset	**Strategy**	**Concat**	**Max**	**Sim-Sum**	**Wtd-Sum**	**Concat**	**Max**	**Sim-Sum**	**Wtd-Sum**
aNLI	OPENBOOK	73.9± 0.8	73.7± 0.1	73.5± 0.7	73.3± 1.0	83.9± 0.5	80.8± 0.9	81.7± 0.6	84.4± 0.4
	REVISION	72.7± 0.3	N/A	N/A	N/A	82.4	N/A	N/A	N/A
	REVISION & OPENBOOK	74.4± 0.2	74.3± 0.1	74.0± 0.9	75.1±0.4	84.2± 0.7	81.4± 0.8	82.6± 0.6	**86.7**± 0.6
PIQA	OPENBOOK	67.8± 0.4	72.4± 0.6	72.6± 1.2	72.5± 0.1	74.8± 0.5	75.2± 0.9	75.6± 0.7	77.1± 0.2
	REVISION	74.5± 0.3	N/A	N/A	N/A	75.2± 0.8	N/A	N/A	N/A
	REVISION & OPENBOOK	67.7± 0.1	73.8± 0.8	76.8± 0.5	76.8± 0.3	75.4± 0.7	76.2± 0.8	76.8± 0.4	**80.2**± 0.6
SIQA	OPENBOOK	70.1± 0.8	67.8± 0.1	70.0± 0.7	70.2± 0.4	76.5± 0.7	77.2± 0.6	77.4± 0.2	78.3± 0.5
	REVISION	69.5± 0.9	N/A	N/A	N/A	76.8± 0.3	N/A	N/A	N/A
	REVISION & OPENBOOK	68.8± 0.4	66.6± 0.4	68.9± 0.1	69.3± 0.6	78.2± 0.3	77.4± 0.9	76.7± 0.5	**79.5**± 0.9

Models/ Accuracy	**aNLI**		**PIQA**		**SIQA**	
	Val	**Test**	**Val**	**Test**	**Val**	**Test**
BERT	67.36	66.75	68.08	69.23	64.88	64.50
GPT-2 XL	N/A	N/A	70.20	69.50	47.50	45.30
RoBERTa	85.05	83.91	76.28	76.80	77.85	76.74
RoBERTa 5 Ensemble	N/A	83.22	N/A	79.66	N/A	78.68
$L2R^2$ [2020]	N/A	**86.81**	N/A	N/A	N/A	N/A
KagNet [2019]	N/A	N/A	N/A	N/A	65.05	64.59
GBR [2020]	N/A	N/A	N/A	N/A	75.64	76.25
UnifiedQA T5 11B [2020]	N/A	80.04	N/A	**89.50**	N/A	**79.75**
Ours: BERT + WS	74.60	74.96	76.82	72.28	70.21	67.22
Ours: RoBERTa + WS	85.90	84.18	80.20	78.24	79.53	78.00



RoBERTA Acc with Increasing Sentences