CLMSM: A Multi-Task Learning Framework for Pre-training on Procedural Text

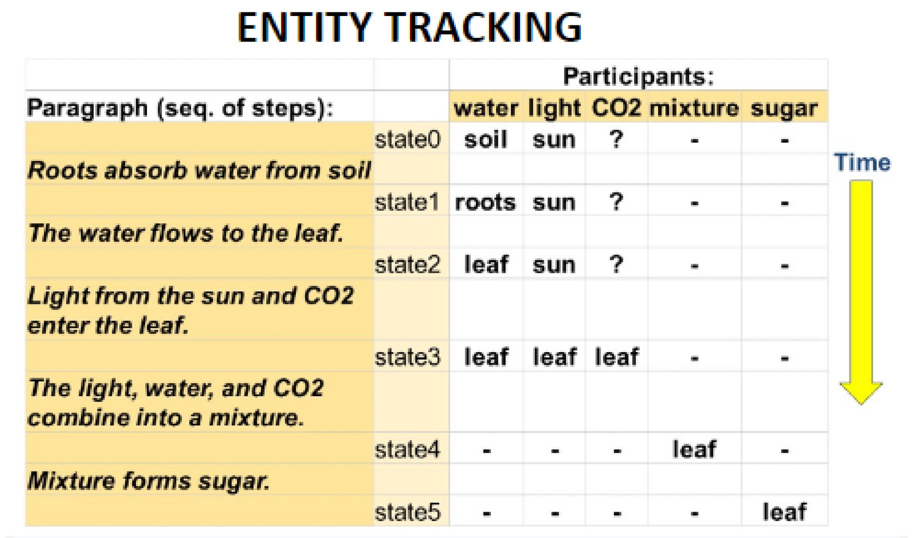
Abhilash Nandy, Manav Nitin Kapadnis, Pawan Goyal, Niloy Ganguly

CNeRG Lab, Indian Institute of Technology Kharagpur, India

EMNLP 2023

Background

- In the realm of procedural text, understanding and processing tasks such as entity tracking and action alignment are pivotal.
- Conventional Pre-training such as MLM does not capture temporal information
- Small size of the procedural reasoning benchmark datasets



Paragraph (seq. of steps):		Participants: water	light	CO2	mixture	sugar
	state0	soil	sun	?	-	-
Roots absorb water from soil						
	state1	roots	sun	?	-	-
The water flows to the leaf.						
	state2	leaf	sun	?	-	-
Light from the sun and CO2 enter the leaf.						
	state3	leaf	leaf	leaf	-	-
The light, water, and CO2 combine into a mixture.						
	state4	-	-	-	leaf	-
Mixture forms sugar.						
	state5	-	-	-	-	leaf

Figure 1: Example Of Entity Tracking Task

Problem Statement

Coming up with Pre-training tasks to solve procedural reasoning tasks.

- Using inter and intra-procedure (stepwise) supervision signals to solve Recipe Alignment and Entity Tracking Tasks
- Fine-tuning of procedural text representation using Contrastive Learning
- Develop understanding of procedural text nature using Mased Step Modelling

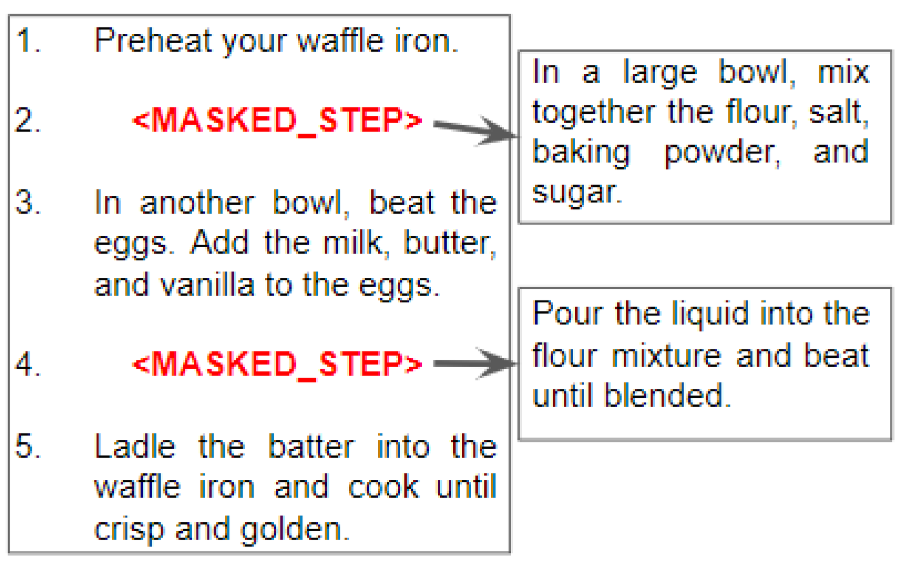


Figure 2: Overview of Masked Step Modelling

Dataset Collection

We collected, and preprocessed a corpus of **2.8 Million** recipes after deduplication was carried out.

- Recipe1M+ dataset (Marin et al., 2018)
- RecipeNLG dataset (Bie n et al., 2020)
- Generating personalized recipes from historical user preferences (Majumder et al., 2019)
- Storyboarding of recipes: Grounded contextual generation. (Chandu et al., 2019)

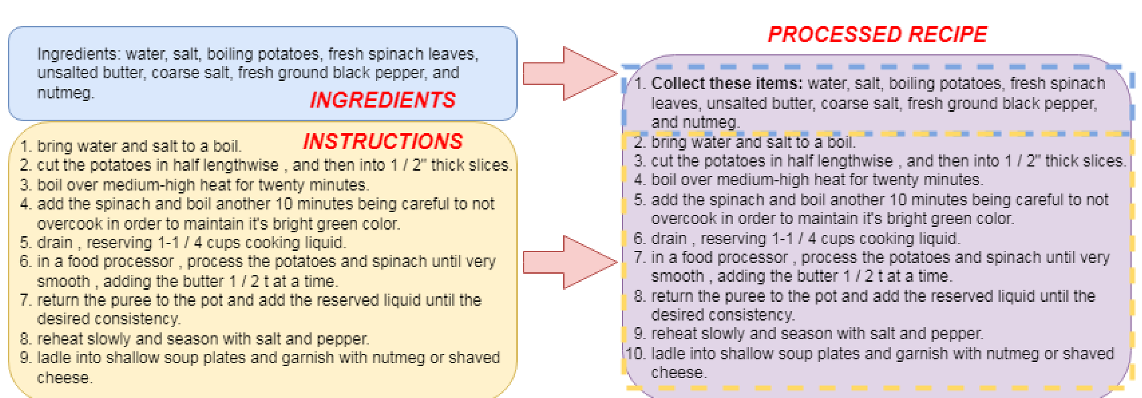


Figure: Modifying a recipe by adding ingredients as a step at the top of the original recipe

Methodology

CLMSM involves using metadata of procedures and a Contrastive Learning strategy to capture global context, while Masked Step Modeling addresses the local context.

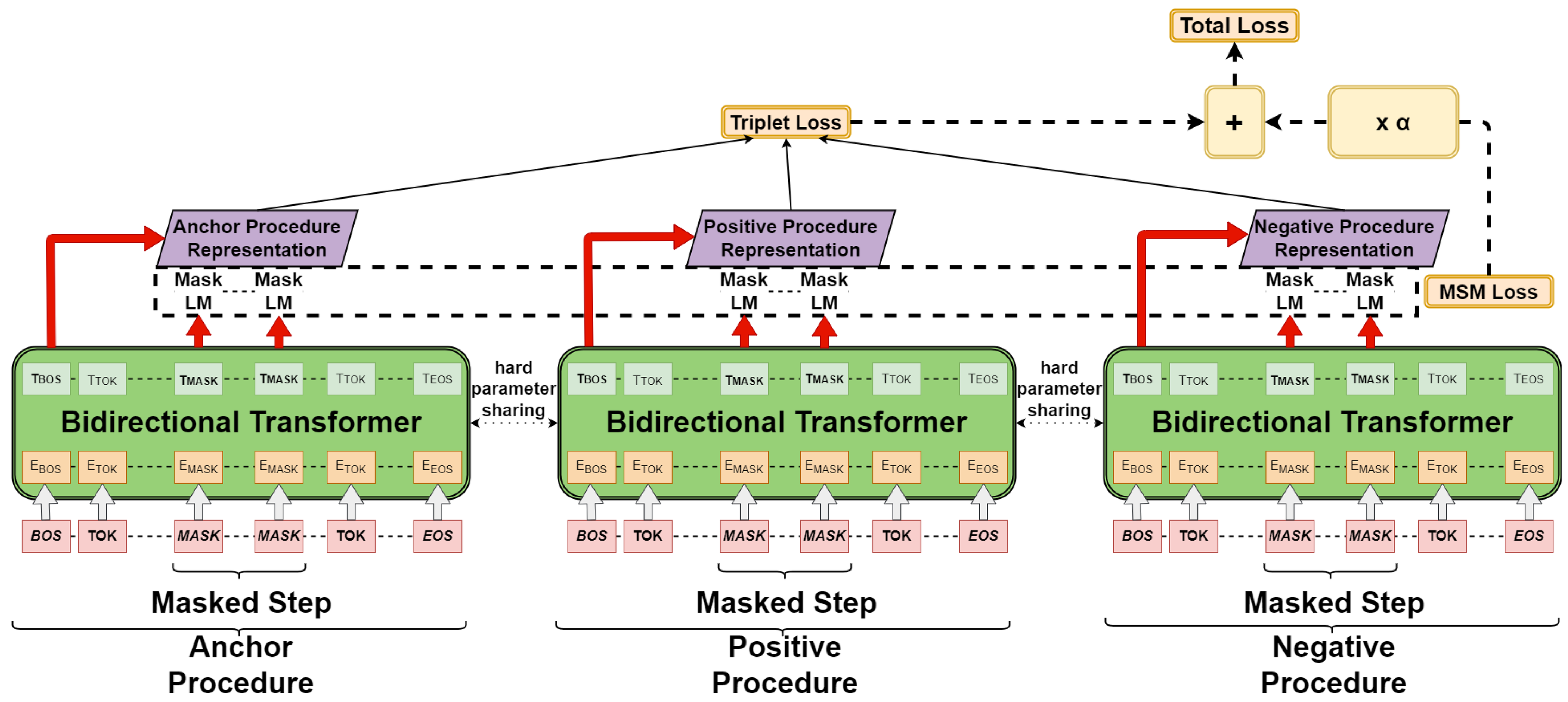


Figure: ***CLMSM*** pre-training framework: Step-Masking and Contrastive Learning objectives are used. MSM Loss obtained by masking step(s) in each recipe is linearly combined with Triplet Loss, and total loss is backpropagated.

Experiments and Results

	TEST ACCURACY
NPN-Model	51.3
KG-MRC	51.6
DYNAPRO	62.9
$\text{RECIPES}_{\text{RB}}$	63.8
$\text{CLMSM}_{RB}(CL)$	65.13
$\text{CLMSM}_{RB}(MSM)$	**65.35**
$\text{CLMSM}_{RB}(CAS.)$	63.99
$\text{CLMSM}_{RB}(RS)$	63.46
$\text{CLMSM}_{RB}(EASY)$	63.68
CLMSM_{RB}	64.62

Table 2: Results when fine-tuned on the NPN-Cooking Dataset - Baselines and CLMSM Variations vs. CLMSM

	TEST CAT. 1 SCORE	TEST CAT. 2 SCORE	TEST CAT. 3 SCORE	AVERAGE SCORE
Rule-based	57.14	20.33	2.4	26.62
Feature-based	58.64	20.82	9.66	29.71
ProLocal	62.7	30.5	10.4	34.53
ProGlobal	63	36.4	35.9	45.1
EntNet	51.6	18.8	7.8	26.07
QRN	52.4	15.5	10.9	26.27
KG-MRC	62.9	40.0	38.2	47.03
NCET	73.7	47.1	41.0	53.93
DYNAPRO	72.4	49.3	**44.5**	55.4
$\text{RECIPES}_{\text{RB}}$	71.33	34.11	30.86	45.43
$\text{RECIPES}_{\text{RL}}$	74.58	47.53	38.37	53.49
$\text{CLMSM}_{RB}(CL)$	67.8	35.62	33.61	45.68
$\text{CLMSM}_{RB}(MSM)$	67.09	36.83	34.03	45.98
$\text{CLMSM}_{RB}(CAS.)$	67.94	36.85	35.94	46.91
$\text{CLMSM}_{RB}(RS)$	65.4	28.57	38.25	44.07
$\text{CLMSM}_{RB}(EASY)$	68.22	35.7	32.1	45.34
CLMSM_{RB}	69.92	39.35	33.89	47.72
CLMSM_{RL}	**77.26**	**54.86**	38.34	**56.82**

Table 3: Results when fine-tuned on the ProPara Dataset - Baselines and CLMSM Variations vs. CLMSM

	TEST CAT. 1 SCORE	TEST CAT. 2 SCORE	TEST CAT. 3 SCORE	AVERAGE SCORE
Falcon-7B (1-shot)	50.42	7.11	4.79	20.77
Falcon-7B (3-shot)	50.85	7.73	5.3	21.29
Falcon-7B-Instruct (1-shot)	50.42	5.42	0.38	18.74
Falcon-7B-Instruct (3-shot)	48.44	3.15	1.94	17.84
GPT-3.5 (1-shot)	53.25	24.66	11.37	29.76
GPT-3.5 (3-shot)	62.43	34.66	15.81	37.63
GPT-4 (1-shot)	57.2	31.08	17.03	35.10
GPT-4 (3-shot)	73.87	**57.7**	26.78	52.78
CLMSM_{RB}	69.92	39.35	33.89	47.72
CLMSM_{RL}	**77.26**	54.86	**38.34**	**56.82**

Table 4: Results on the ProPara Dataset - LLMs vs. CLMSM

Results on Entity Tracking Tasks

	TEST ACCURACY
$\text{RECIPES}_{\text{BERT}}$	64.22
$\text{CLMSM}_{BERT}(CL)$	55.96
$\text{CLMSM}_{BERT}(MSM)$	61.47
$\text{CLMSM}_{BERT}(CAS.)$	55.96
$\text{CLMSM}_{BERT}(RS)$	54.31
$\text{CLMSM}_{BERT}(EASY)$	57.06
CLMSM_{BERT}	**66.97**

Table 6: Results when fine-tuned on the ARA Dataset - Baselines vs. CLMSM

Results on Action Alignment across procedures

Qualitative Analysis

- ***CLMSM*** outperforms various types of baselines **(non-neural, LSTM-based, Knowledge Graph-based, transformer-based)**
- **In-domain pre-training helps in learning temporal entity-based information** required for entity tracking. This also shows the **out-of-domain adaptation** of CLMSM.

Key Insights

- ***CLMSM***, a novel pre-training framework specialized for procedural data, that leverages metadata information of procedures and devises a Contrastive Learning strategy to capture global context, and Masked Step Modeling to capture local context.
- ***CLMSM*** performs better than best baseline for tracking entities whose locations are implicit in nature, i.e., the location at current step requires retention of knowledge from step far in past.
- Even though ***CLMSM*** is pre-trained on recipes, it shows improvement on an open-domain procedural task on ProPara Dataset, thus showing its generalizability.

Paper Link

Code Base Link

For Further Information

Arxiv: `https://arxiv.org/abs/2310.14326`