Distilling Script Knowledge from Large Language Models for Constrained Language Planning

Siyu Yuan[1] (syyuan21@m.fudan.edu.cn), Jiangjie Chen[1], Ziquan Fu[2], Xuyang Ge[1],
Soham Shah[3], Charles Robert Jankowski[3], Yanghua Xiao[1], Deqing Yang[1]

[1]Fudan University, [2]System Inc, [3]Brain Technologies, Inc.

Constrained Language Planning

🥞 How to make a **strawberry** cake ✖ Previous work focus on planning for **abstract goals** of stereotypical activities

🧁 How to make a **chocolate** cake ☑ We exploit **specific goals** with multi-facet **constraints**.
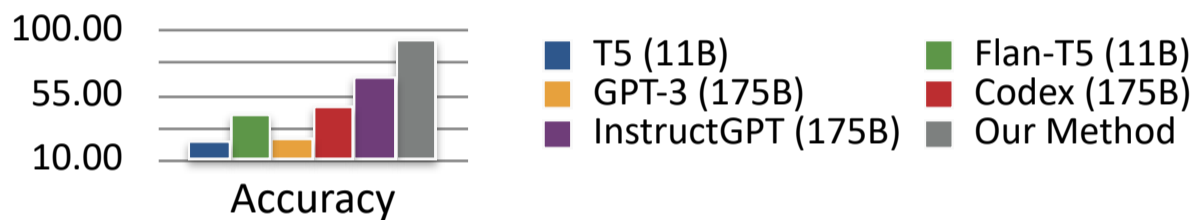
Contribution

❖ Establish the **constrained language planning** problem

❖ Evaluate the **constrained language planning ability** of LLMs and develop an **over-generate-then-filter** method.

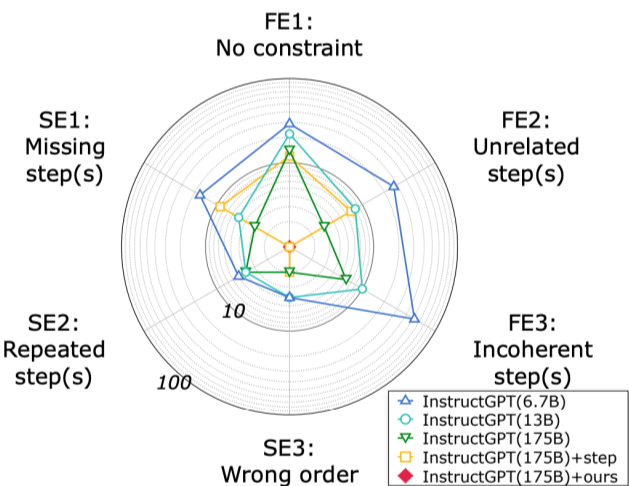❖ Use LLMs to generate a dataset **(CoScript)** for constrained language planning

Analysis of Large Language Models

Dataset: wikiHow+Generated Constraints
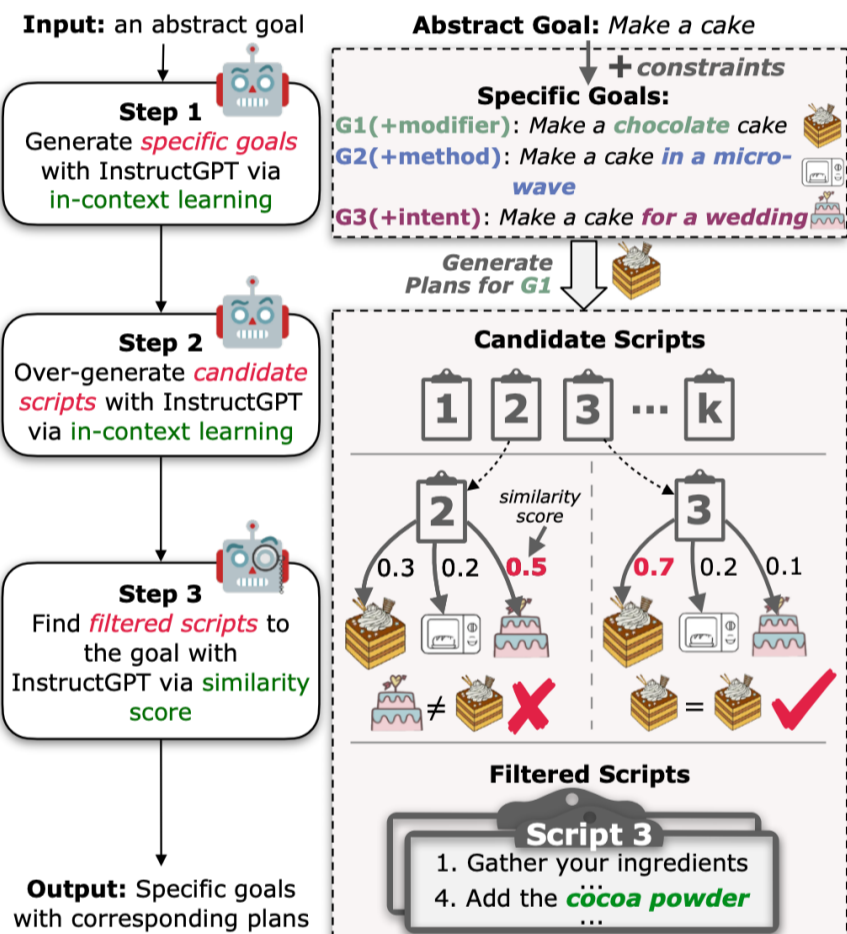
LLMs can not achieve satisfactory performance

- T5 (11B)
- GPT-3 (175B)
- InstructGPT (175B)
- Flan-T5 (11B)
- Codex (175B)
- Our Method

Accuracy

LLMs sometimes do not plan faithfully to the constraints.

FE1: No constraint
FE2: Unrelated step(s)
FE3: Incoherent step(s)
SE1: Missing step(s)
SE2: Repeated step(s)
SE3: Wrong order

- InstructGPT(6.7B)
- InstructGPT(13B)
- InstructGPT(175B)
- InstructGPT(175B)+step
- InstructGPT(175B)+ours

LLMs varies considerably for goals of different categories

	6.7B	13B	175B	Step	Ours
Work	0.0	0.67	1.0	1.0	1.0
Sports	0.25	0.5	1.0	1.0	1.0
Relationships	0.0	0.5	1.0	1.0	1.0
Personal Care	0.14	0.33	0.5	0.6	1.0
Home	0.17	0.3	0.5	0.6	1.0
Traditions	0.25	0.4	0.5	0.57	1.0
Hobbies	0.1	0.42	0.6	0.67	0.88
Health	0.33	0.5	0.57	0.75	0.9
Food	0.28	0.66	0.78	0.78	0.94
Business	0.33	0.5	1.0	0.67	1.0
Family Life	0.0	0.0	1.0	0.67	1.0
Education	0.2	0.33	0.5	0.67	1.0
Electronics	0.25	0.38	0.6	0.62	0.92
Vehicles	0.0	0.0	0.75	0.5	1.0
Arts	0.33	0.29	1.0	0.5	1.0

Our method

Input: an abstract goal

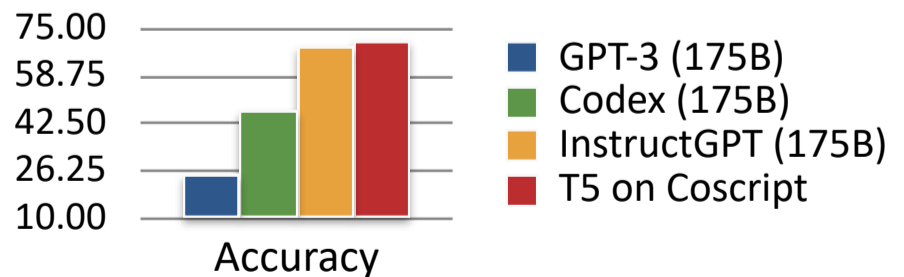**Step 1** Generate *specific goals* with InstructGPT via **in-context learning**

Step 2 Over-generate *candidate scripts* with InstructGPT via **in-context learning**

Step 3 Find *filtered scripts* to the goal with InstructGPT via **similarity score**

Output: Specific goals with corresponding plans

Abstract Goal: *Make a cake*
+ constraints
Specific Goals:
G1(+modifier): Make a **chocolate** cake
G2(+method): Make a cake **in a microwave**
G3(+intent): Make a cake **for a wedding**

Generate Plans for G1

Candidate Scripts
1 2 3 ... k

similarity score
0.3 0.2 **0.5**   **0.7** 0.2 0.1

≠ ✗   = ✓

Filtered Scripts

Script 3
1. Gather your ingredients
4. Add the **cocoa powder**
....

Coscript for Smaller Language Models

Motivation

☑ To enable constrained language planning ability for smaller models.

Method

☑ Follow the idea of **symbolic knowledge distillation**

☑ Generated 55,000 Scripts with constraint from LLMs based on our method → **Coscript Dataset**

☑ Humans annotate validation and test set.

Data Statistics

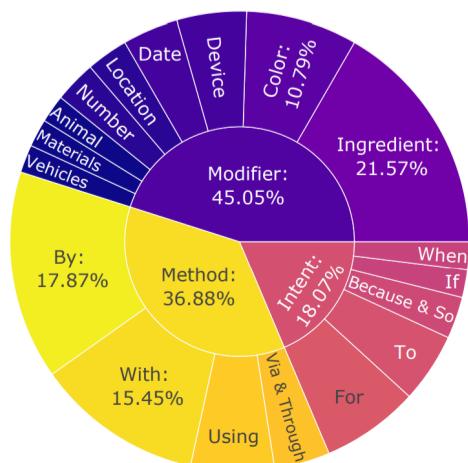| Dataset	# Size	# UT	Avg$_{\mathcal{G}_c}$ #	Avg$_S$ #
proScript	6,414	8,826	0	5.45
wikiHow	112,111	158,117	0.42	5.93
CoScript	55,000	76,317	4.13	5.96

Constraint Distribution

Modifier: 45.05%
Ingredient: 21.57%
Method: 36.88%
By: 17.87%
With: 15.45%
Intent: 18.07%
Color: 10.79%
Date, Device, Number, Location, Animal, Materials, Vehicles
Using, When, If, Because & So, To, For, Via & Through

Smaller LMs fine-tuned on CoScript can generate higher quality scripts than LLMs

- GPT-3 (175B)
- Codex (175B)
- InstructGPT (175B)
- T5 on Coscript

Accuracy

CoScript consistently improves model performance in constraint faithfulness

- T5-Large
- GPT-2-Large

α: % of CoScript