AI Chains: Transparent and Controllable Human-AI Interaction by Chaining LLM Prompts

Tongshuang (Sherry) Wu  @tongshuangwu
University of Washington

Michael Terry
Carrie J. Cai  @Carryveggies
Google Research
Large Language Models (LLMs) are helpful.

Respond to natural language prompts (instructions), accessible to end users!

---

**Translation**

**English sentence**

Where’s the bus station?

**Prompt**

Given the English sentence, translate to French.

**English:** Where’s the bus station?

**French:** Où est la gare routière?

---

*Brown, Tom B., et al. "Language models are few-shot learners." arXiv 2020*
LLM may **struggle with complex tasks.**

*Example: Peer review rewriting*

---

**Rewriting**

**Original feedback**

Sherry could improve her presentation...

**More friendly feedback**

Sherry, you have too many words on your...

---

Given the following feedback, rewrite it into a friendly paragraph with concrete suggestions for each of Alex’s presentation problems.

**Original feedback:**

Sherry could improve her presentation skills. She has too much text on her slides. Her presentation meanders from topic to topic without a clear structure. She also does not engage with her audience when she presents.

**More friendly feedback:**

Sherry, you have too many words on your slides. You should use images and bullet points to help get your message across. You should have a clear structure for the presentation. You should also engage with your audience.
LLM may struggle with complex tasks.

Example: Peer review rewriting

Model structure has limitations: reasoning capabilities, exposure bias, etc.

...that end users don’t know or understand.

How can we help end users interpret and improve model outputs, when they cannot understand or change the model?
LLM may struggle with complex tasks.

Example: Peer review rewriting

How can we help end users interpret and improve model outputs, when they cannot understand or change the model?

AI Chain ⚛: Break down an original task into sub-tasks, to provide explicit knobs to users on sub-tasks.

More friendly feedback

Sherry, you have too many words on your slides. You should use images and bullet points to help get your message across. You should have a clear structure for the presentation. You should also engage with your audience.
The input-output mapping is convoluted.

Given the following feedback, rewrite it into a friendly paragraph with concrete suggestions for each of Alex’s presentation problems.

Original feedback:
Sherry could improve her presentation skills. She has too much text on her slides. Her presentation meanders from topic to topic without a clear structure. She also does not engage with her audience when she presents.

More friendly feedback:
Sherry, you have too many words on your slides. You should use images and bullet points to help get your message across. You should have a clear structure for the presentation. You should also engage with your audience.
Small tasks are more interpretable and controllable.

Given the Presentation problem, the following is a list of improvement suggestions.

**Problem:** Does not engage

**Suggestions for improvements:**

1) Ask the audience questions
2) Walk around the room

Use humor
The fix can be **propagated** to related sub-tasks!

Write one friendly paragraph to cover all the suggestions.

Suggestions:
1) **Ask the audience questions**
2) **Use humor**

Paragraph
You may want to ask more questions to engage the audience. Humor always helps!

---

Chain =
Independent LLM runs per sub-task
+ inter-step transformation

Presentation problem
Does not engage

Create ideas

Suggestions for improvements
1) Ask the audience questions
2) Walk around the room
Use humor

Compose points

Paragraph
You may want to ask more questions to engage...
Chain =

Independent LLM runs per sub-task

1. Identify all presentation problems
2. Ideate suggestions per problem
3. Compose them into a paragraph

+ inter-step transformation

Suggestions for improvements

1) Ask the audience questions
2) Walk around the room

Use humor

Presentation problem

Does not engage

Compose points

Paragraph

You may want to ask more questions to engage...
A three-step Chain for peer review rewriting

1. Split all presentation problems

Given the Presentation problem, the following is a list of suggestions.

Feedback:
Sherry could improve her presentation skills. She has too much text on her slides. Her presentation meanders from topic to topic without a clear structure. She also does not engage with her audience when she presents.

Presentation problems:
1) Too much text on slides
2) Meanders from topic to topic
Given the Presentation problem, the following is a list of suggestions.

**Problem**: Meanders from topic to topic

**Suggestions for improvements**:
1) Split the talk into sections
2) Read outlines and check off each section as you go

**Problem**: Too much text

**Suggestions for improvements**:
1) Use more graphics
2) Use bullet points

A three-step Chain for peer review rewriting

**Ideate suggestions per problem**

Given the Presentation problem, the following is a list of suggestions.

**Problem**: Meanders from topic to topic

**Suggestions for improvements**:
1) Split the talk into sections
2) Read outlines and check off each section as you go
Given the Presentation problem, the following is a list of suggestions.

**Problem:** Meanders from topic to topic

**Suggestions for improvements:**
1) Split the talk into sections
2) Read outlines and check off each section as you go

**Problem:** Too much text

**Suggestions for improvements:**
1) Use more graphics
2) Use bullet points

**Problem:** Meanders from topic to topic

**Suggestions for improvements:**
1) Split the talk into sections
2) Read outlines and check off each section as you go
Given the Presentation problem, the following is a list of suggestions.

**Problem**: Meanders from topic to topic

**Suggestions for improvements**:
1) Split the talk into sections
2) Read outlines and checkout each section as you go.

**Problem**: Too much text

**Suggestions for improvements**:
1) Use more graphics
2) Use bullet points
Write one friendly paragraph that covers all the presentation problems and suggestions.

**Presentation problems:**
1) Too much text on slides
2) Meanders from topic to topic
3) Does not engage with audience when presenting.

**Short suggestions for improvements:**
1) Use more graphics
2) Use bullet points
3) Split the talk into sections
4) Read outlines and check off each section as you go
5) Ask the audience questions

Friendly paragraph:
Sherry, your presentation was interesting! However, I noticed that you have a lot of information on your slides. It might be helpful to vary pictures with text so that it is easier to follow. Also, you might consider the flow of your theme. If it were me, I would have divided it into three sections and then used your conclusion. You may also want to add some humor, and ask more questions to engage the audience.
Chain =
Independent LLM runs per sub-task

1. Identify all presentation problems
2. Ideate suggestions per problem
3. Compose them into a paragraph

+ inter-step transformation

Less is more!

Sherry, your presentation was interesting! However, I noticed that you have a lot of information on your slides. It might be helpful to vary pictures with text so that it is easier to follow.
Chaining improves human-AI interactions

**Transparent**: better convey model’s goal per step

“Chaining helped you speak the language. It lift[ed] up the hood and showed what’s happening at different phrases.”

**Rating Metrics**

- **Transparent**
- **Controllable**
- **Collaborative**

**Controllable**: Explicit “knobs”

“Chaining makes it easier to fine-tune things… Too much freedom [without Chaining] is a curse.”
Effective decomposition: LLM Challenge vs. Ability

LLM face challenges, when we ask for too much (literature):

Operations for LLM Chaining
Nudge users towards scoped inputs/outputs.
Effective decomposition: LLM Challenge vs. Ability

LLM face challenges, when we ask for too much (literature):

Generate too much content.
Exposure bias, overfit to their own generations.
e.g., produce redundant content.

Operations for LLM Chaining
Nudge users towards scoped inputs/outputs.

Gather additional information from LLMs
e.g. Ideation

Chat topics
Small talk
Transportation
Accommodation

Examples in English
Do you like the weather?
Where’s the bus station?
How do I go to the Louvre?
I will check out at noon.

Examples in French
Vous aimez le temps?
Où est la gare routière?
Comment aller au Louvre?
je partirai à midi.

City to visit
Paris
Effective decomposition: LLM Challenge vs. Ability

LLM face challenges, when we ask for too much (literature):

- **Generate too much content.**
- **Handle too specific data formats.** Confused by mismatched text distribution e.g., Handled natural text better than JSON format.

---

**Operations for LLM Chaining**

Nudge users towards scoped inputs/outputs.

- **Gather additional information from LLMs**
- **Re-organize the input** e.g. Rewriting

---

**Original VegaLite Spec**

```
"size": {
  "field": "Origin",
  "type": "nominal"
}
```

**Viz. description**

A) [encoding: size] has [data type: nominal]

B)...

**Fixed VegaLite Spec**

```
"size-color": {
  "field": "Origin",
  "type": "nominal"
}
```

---

**Rewriting**

- For **Encoding**
- For **Data Type**
Effective decomposition: LLM Challenge vs. Ability

LLM face challenges, when we ask for too much (literature):

- Generate too much content.
- Handle too specific data type.
- Resolve too many sub-parts.
  Not enough reasoning capability.
  e.g., if-this-then-that branching logic

Operations for LLM Chaining
Nudge users towards scoped inputs/outputs.

- Gather additional information from LLMs
- Re-organize the input
- Validate and categorize the input
  e.g. Classification

Operations for LLM Chaining

<table>
<thead>
<tr>
<th>Input</th>
<th>Let's go</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Let's get</td>
</tr>
</tbody>
</table>

Rewriting

Shorthand

Phrase

Generation

<table>
<thead>
<tr>
<th>Input</th>
<th>Let's go</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Let's go check it out</td>
</tr>
<tr>
<td></td>
<td>Let's go to the mall!</td>
</tr>
</tbody>
</table>
Effective decomposition: LLM Challenge vs. Ability

LLM face challenges, when we ask for too much (literature):

- Exposure bias.
- Sub-optimal data format.
- Lack multi-step reasoning capability

Operations for LLM Chaining
Nudge users towards scoped inputs/outputs.

- Gather additional information from LLMs
- Re-organize the input
- Validate and categorize the input

AI Chains *raises the ceiling* for what people can use LLMs.
Chaining for prototyping AI-infused applications

Music play
I will play the following:
1. Get back
2. Hey Jude
3. Love me Do!

Who’re some Country artists?
Garth Brooks, George Strait, ...

Hey! what up?
I’m chillin‘, how can I help you?

PromptChainer: Chaining Large Language Model Prompts through Visual Programming
Tongshuang Wu, Ellen Jiang, Aaron Donsbach, Jeff Gray, Alejandra Molina, Michael Terry, Carrie Cai
CHI LBW 2022

AI Chains lower the floor for end-users prototyping with LLMs.
**AI Chain: The Big Picture**

**AI Chain**: Break down an original task into sub-tasks, to provide explicit knobs to users on sub-tasks.

Raise the ceiling of what human-LLMs can do, Transparent and controllable interactions, lower the floor for end-users prototyping with LLMs. Gains purely from novel interaction, no model retraining.

- Paper
- LBW follow-up
- Video demo