

CHI 2022

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

Tongshuang (Sherry) Wu @tongshuangwu  
University of Washington

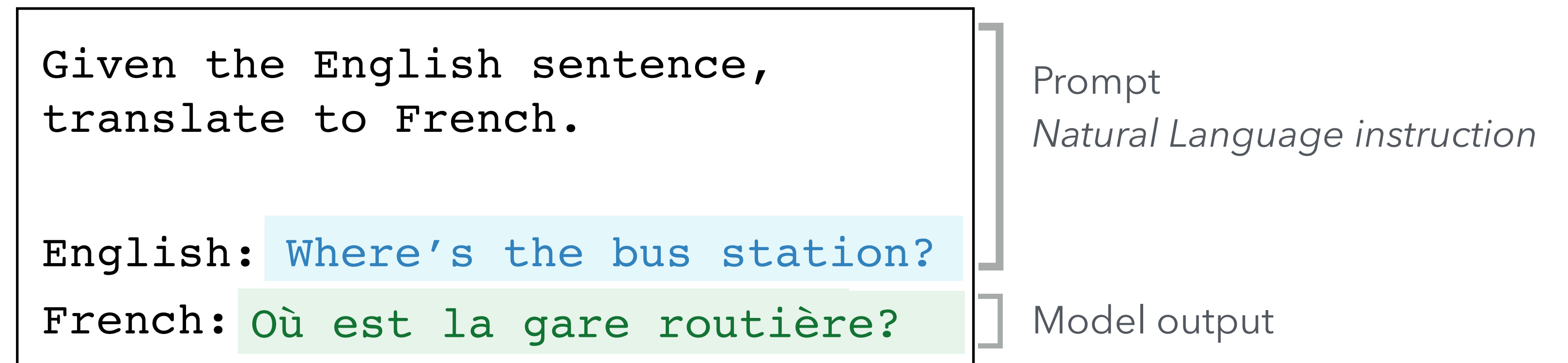
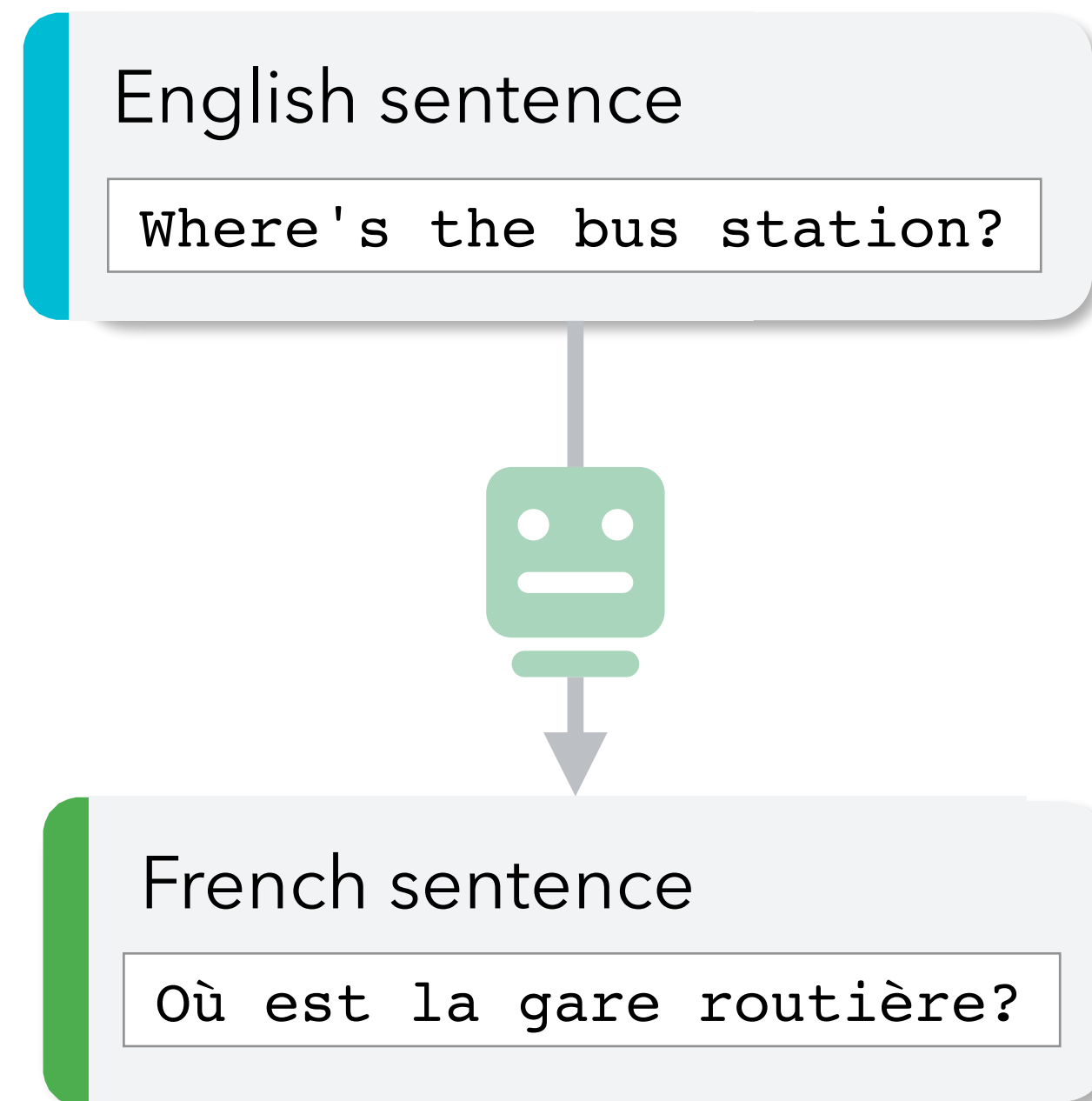
Michael Terry  
Carrie J. Cai @Carryveggies  
Google Research



# Large Language Models (LLMs) are **helpful**.

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

## *Translation*



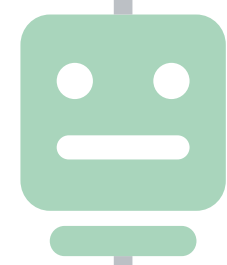
# LLM may struggle with complex tasks.

*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.

Sherry could improve her presentation...

Original feedback:

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

More friendly feedback

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

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*

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.**

Original feedback:

Sherry could improve her presentation skills. She has too

More friendly feedback

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

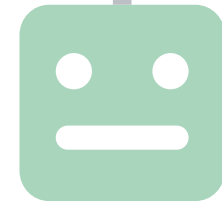
Sherry, you have too many words on your slides. You should use images and bullet points to help get your message across. You should have a clear structure for the presentation. You should also engage with your audience.

# The input-output mapping is **convoluted**.

## *Rewriting*

### Original feedback

Sherry could improve her presentation...



**Confusing mapping!**

### 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.

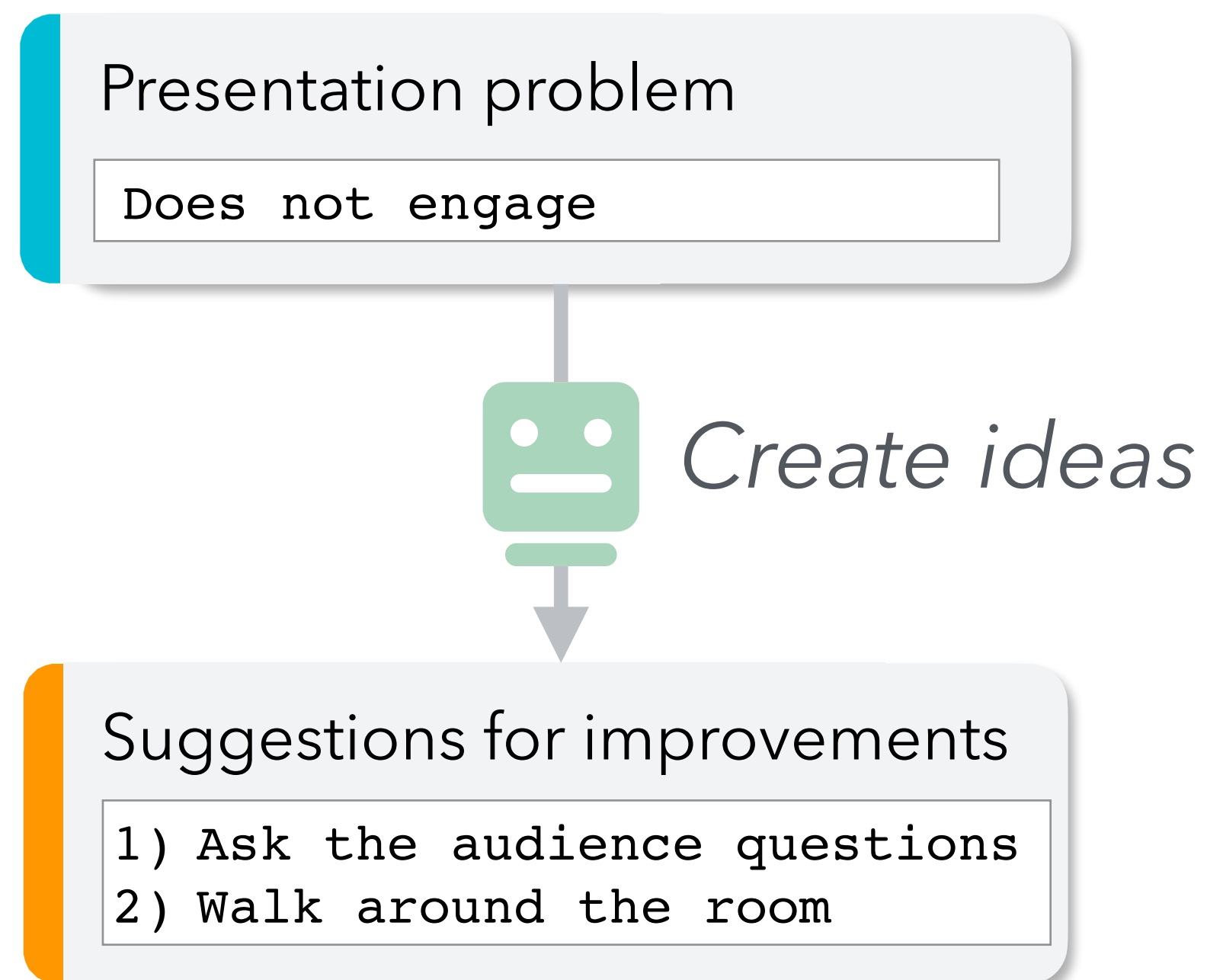
#### Original feedback:

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.

# Small tasks are more **interpretable** and **controllable**.



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

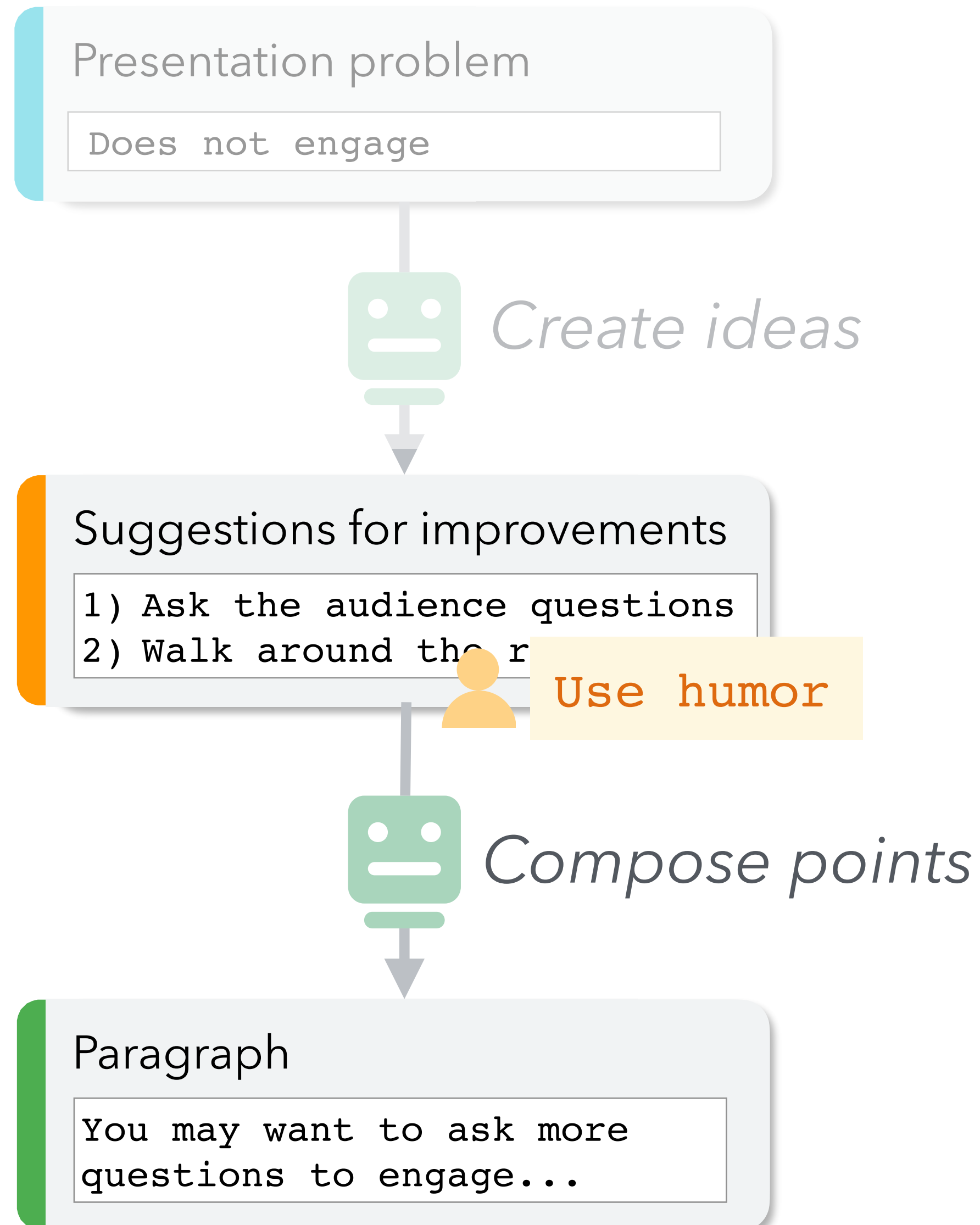
Problem: Does not engage

Suggestions for improvements:

1) Ask the audience questions

 Use humor

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



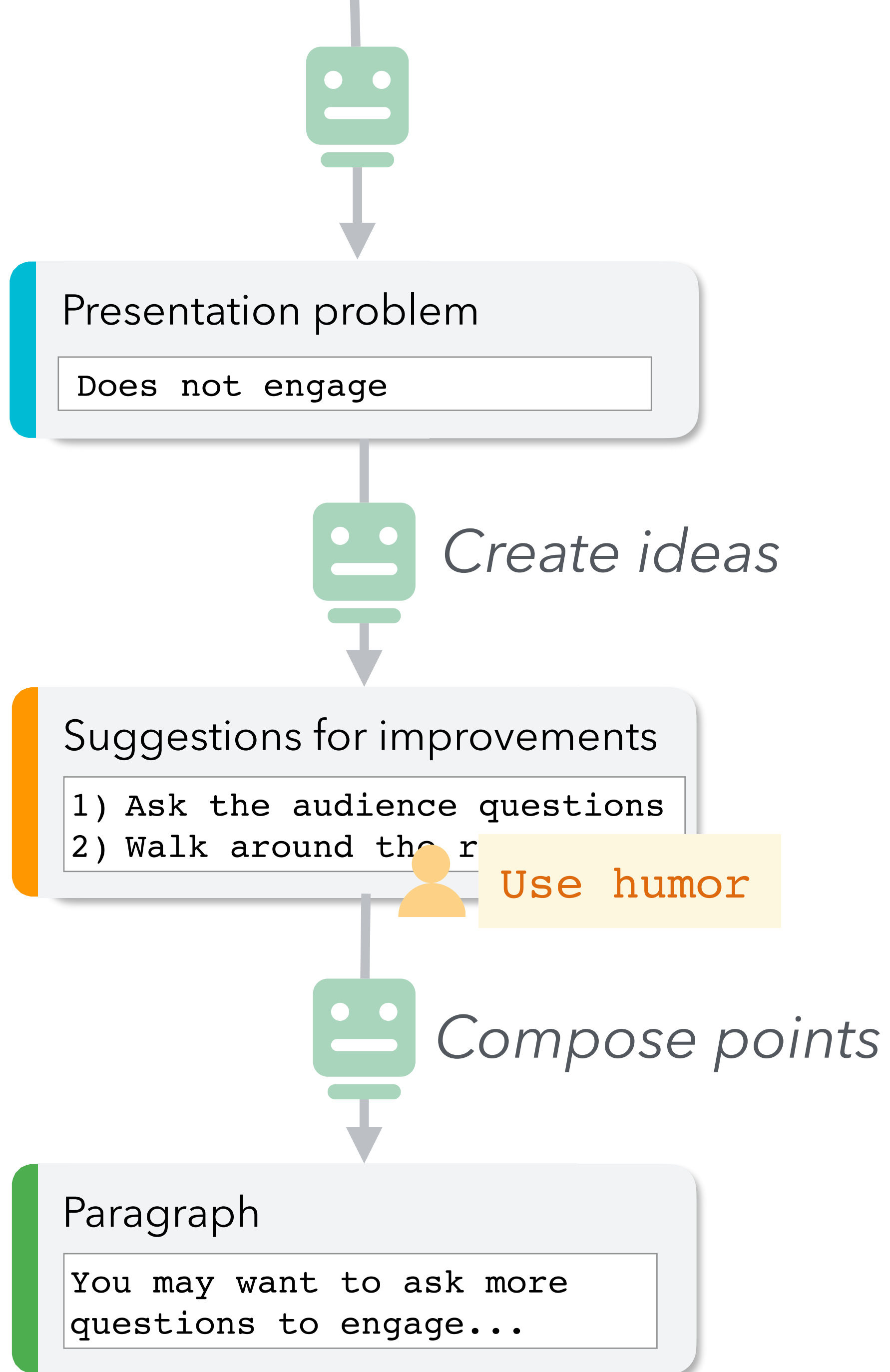
Write one friendly paragraph to cover all the suggestions.

Suggestions:

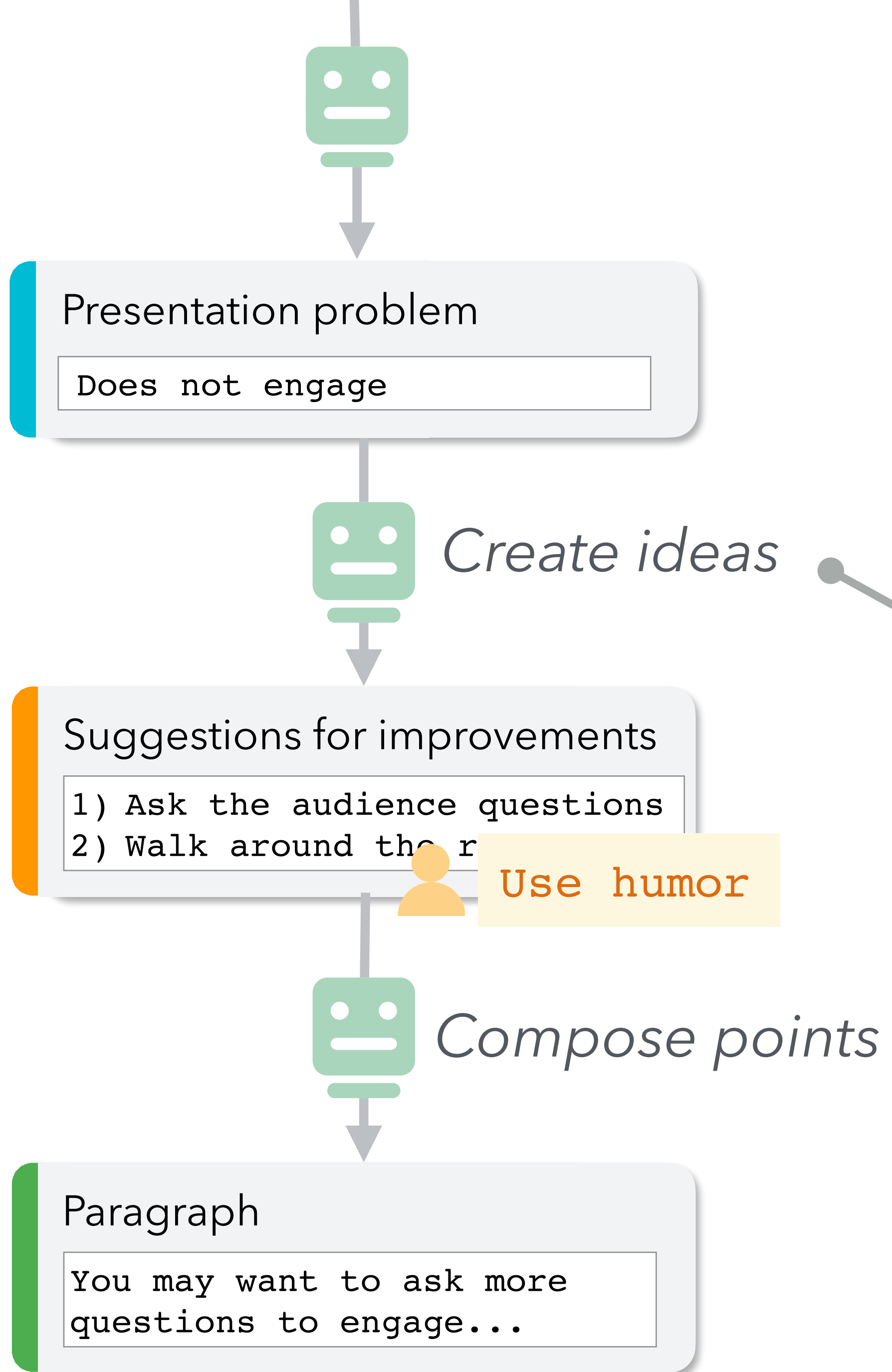
- 1) Ask the audience questions
- 2) Use humor

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

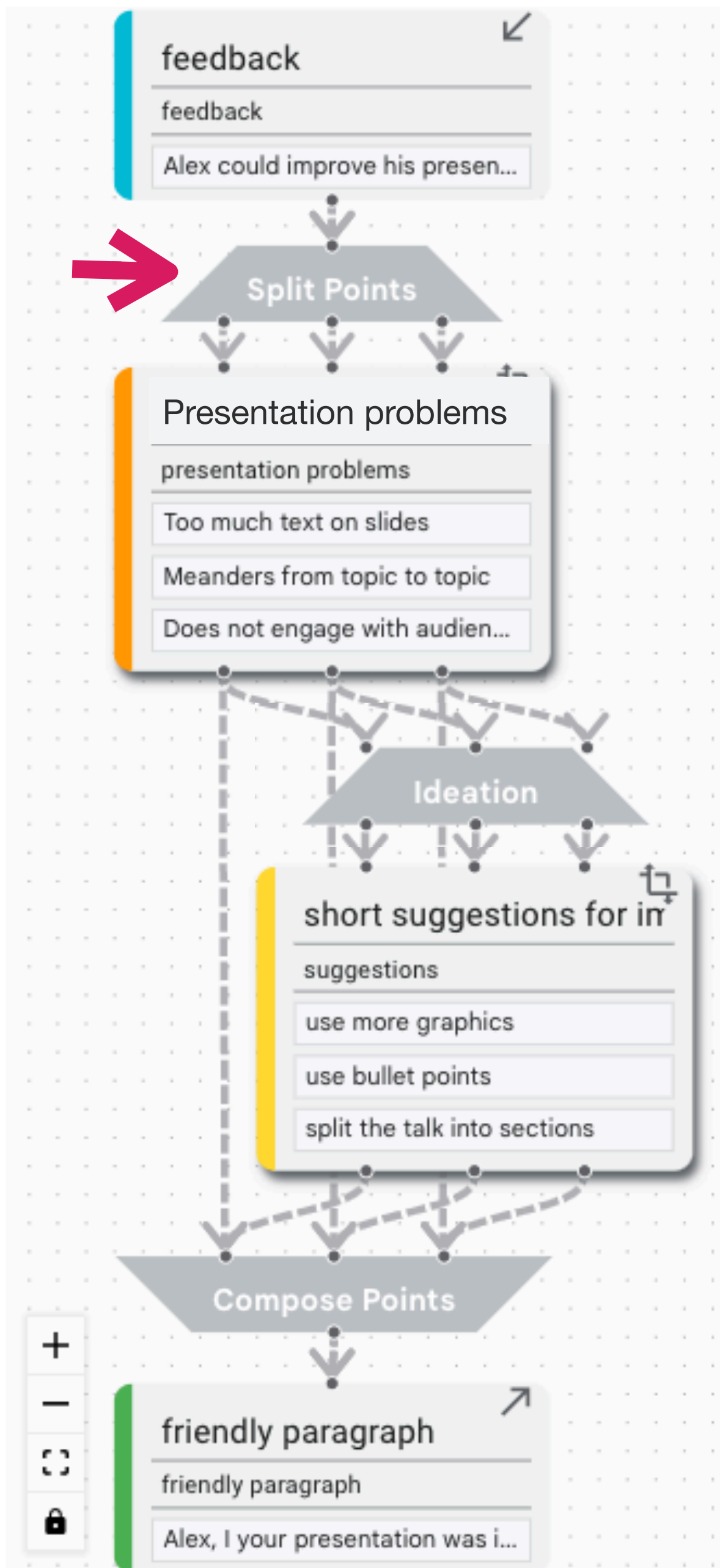


# Chain =

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



[[Instruction block]]

Split the given feedback into a list of presentation problems.  
[[Append running blocks]]

[[Running blocks]]

[[Prepend instruction blocks]]  
Feedback :  
Sherry could improve his presentation skills. She has too much text on his slides. Her presentation meanders from topic to topic without a clear structure. She also does not engage with her audience when she presents.  
Alex's presentation problems :  
1) Too much text on slides -  
2) Meanders from topic to topic -  
3) Does not engage with audience when presenting. -  
+

# A three-step Chain for peer review rewriting

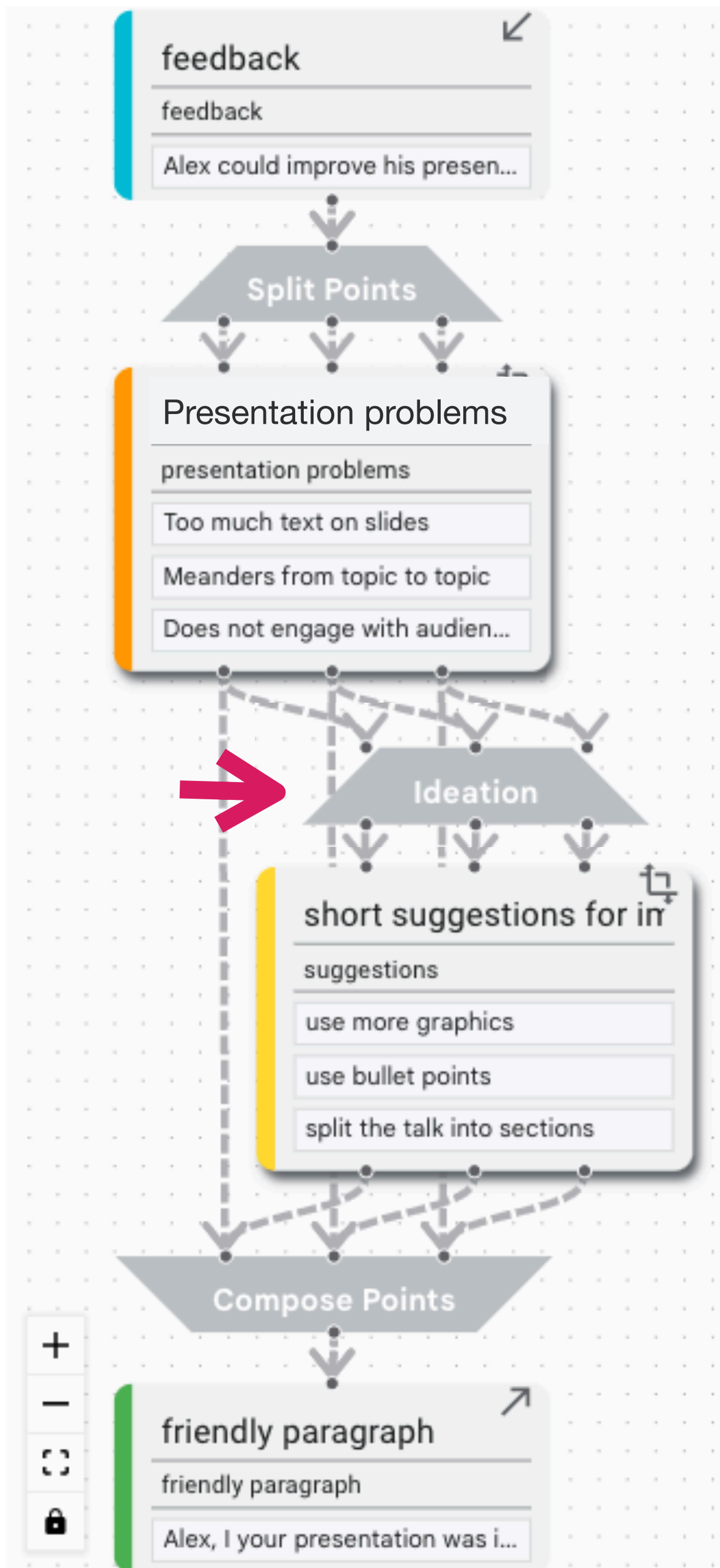
1

**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. Her also does not engage with her audience when she presents.

**Presentation problems:**  
1) Too much text on slides  
2) Meanders from topic to topic



```
[[Instruction block]]
[[Append running blocks]]
Given presentation problems, the following is a list of suggestions.
```

```
[[Running blocks]]
[[Prepend instruction blocks]]
presentation problems :
short suggestions for improvement :
1) use more graphics
2) use bullet points
+
Given presentation problems, the following is a list of suggestions.
Alex's presentation problems:
Too much text on slides
short suggestions for improvement:
1) use more graphics
2) use bullet points
3)
Continue Overwrite prev results Run Remove As example
```

```
[[Prepend instruction blocks]]
presentation problems : : Meanders from topic to topic
short suggestions for improvement :
1) split the talk into sections -
2) read outlines and check off each section as you go -
+
Continue Overwrite prev results Run Remove As example
```

# A three-step Chain for peer review rewriting

## 2

### Ideate suggestions per problem

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

**Problem:** Too much text

**Suggestions for improvements:**

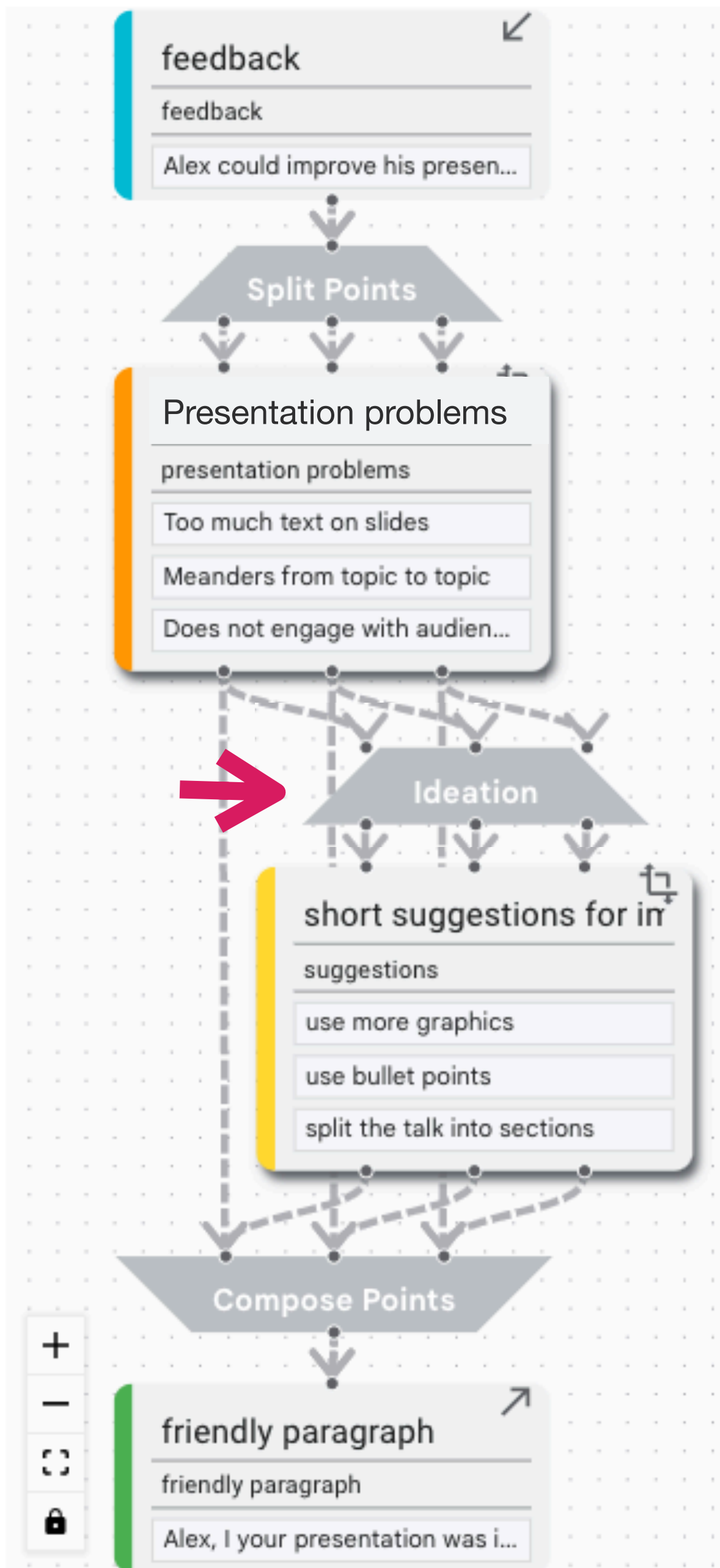
- 1) Use more graphics
- 2) Use bullet points

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.



[[Instruction block]]

Given presentation problems, the following is a list of suggestions.  
[[Append running blocks]]

[[Running blocks]]

[[Prepend instruction blocks]]  
presentation problems :  
short suggestions for improvement :

- 1) use more graphics
- 2) use bullet points

+

Given presentation problems, the following is a list of suggestions.

Alex's presentation problems:  
Too much text on slides  
short suggestions for improvement: 1) use more graphics  
2) use bullet points  
3)

Continue  Overwrite prev results Run Remove As example

[[Prepend instruction blocks]]  
presentation problems : : Meanders from topic to topic  
short suggestions for improvement :

- 1) split the talk into sections -
- 2) read outlines and check off each section as you go -

+

Continue  Overwrite prev results Run Remove As example

# A three-step Chain for peer review rewriting

2

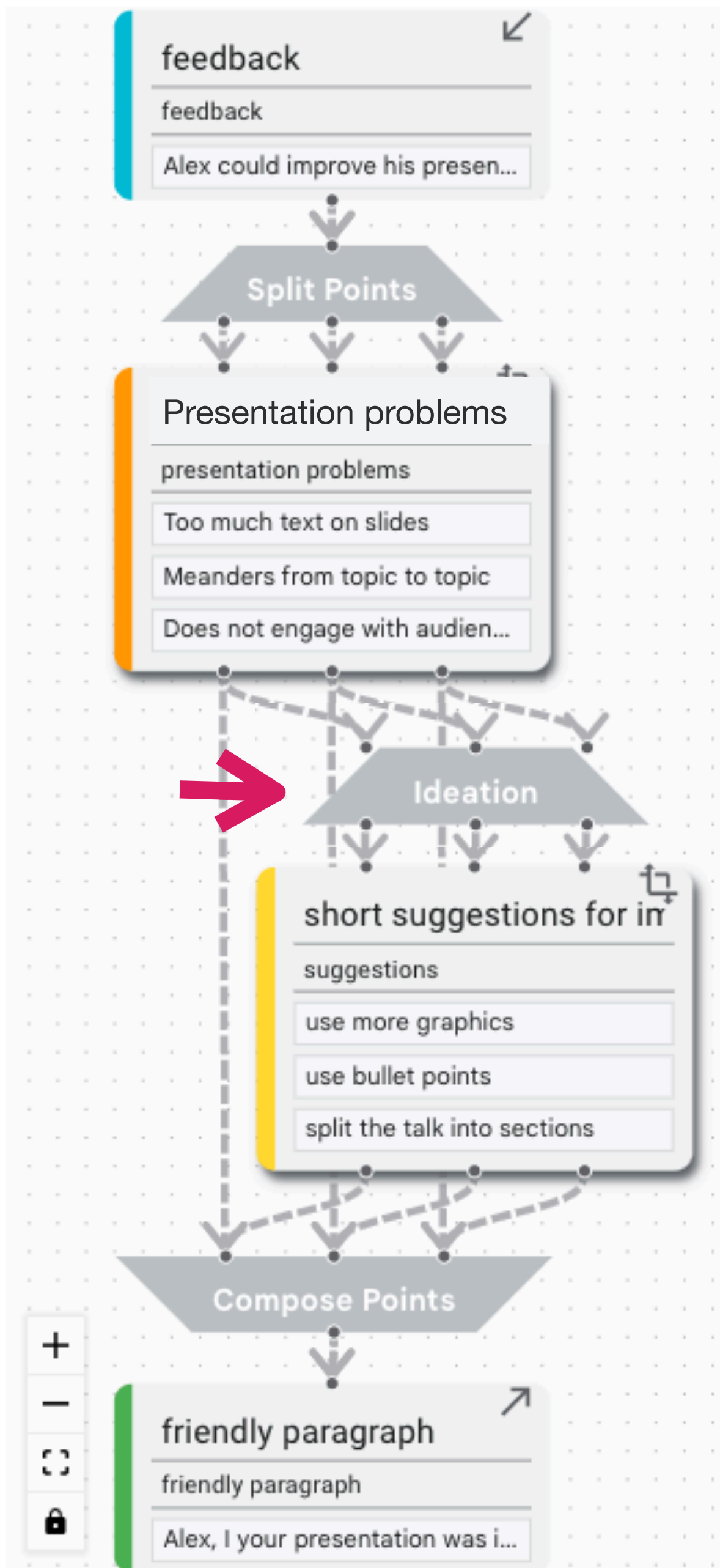
Ideate *suggestions per problem*

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

**Problem:** Too much text  
**Suggestions for improvements:**  
1) Use more graphics  
2) Use bullet points

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.



[[Instruction block]]

Given presentation problems, the following is a list of suggestions.  
[[Append running blocks]]

[[Running blocks]]

[[Prepend instruction blocks]]  
presentation problems :  
short suggestions for improvement :

1) use more graphics  
2) use bullet points

Given presentation problems, the following is a list of suggestions.

Alex's presentation problems:  
Too much text on slides  
short suggestions for improvement: 1) use more graphics  
2) use bullet points  
3)

Continue Overwrite prev results Run Remove As example

[[Prepend instruction blocks]]  
presentation problems : : Meanders from topic to topic  
short suggestions for improvement :

1) split the talk into sections  
2) read outlines and check off each section as you go

Continue Overwrite prev results Run Remove As example

# A three-step Chain for peer review rewriting

2

Ideate *suggestions per problem*

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

**Problem:** Too much text  
**Suggestions for improvements:**  
1) Use more graphics  
2) Use bullet points

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

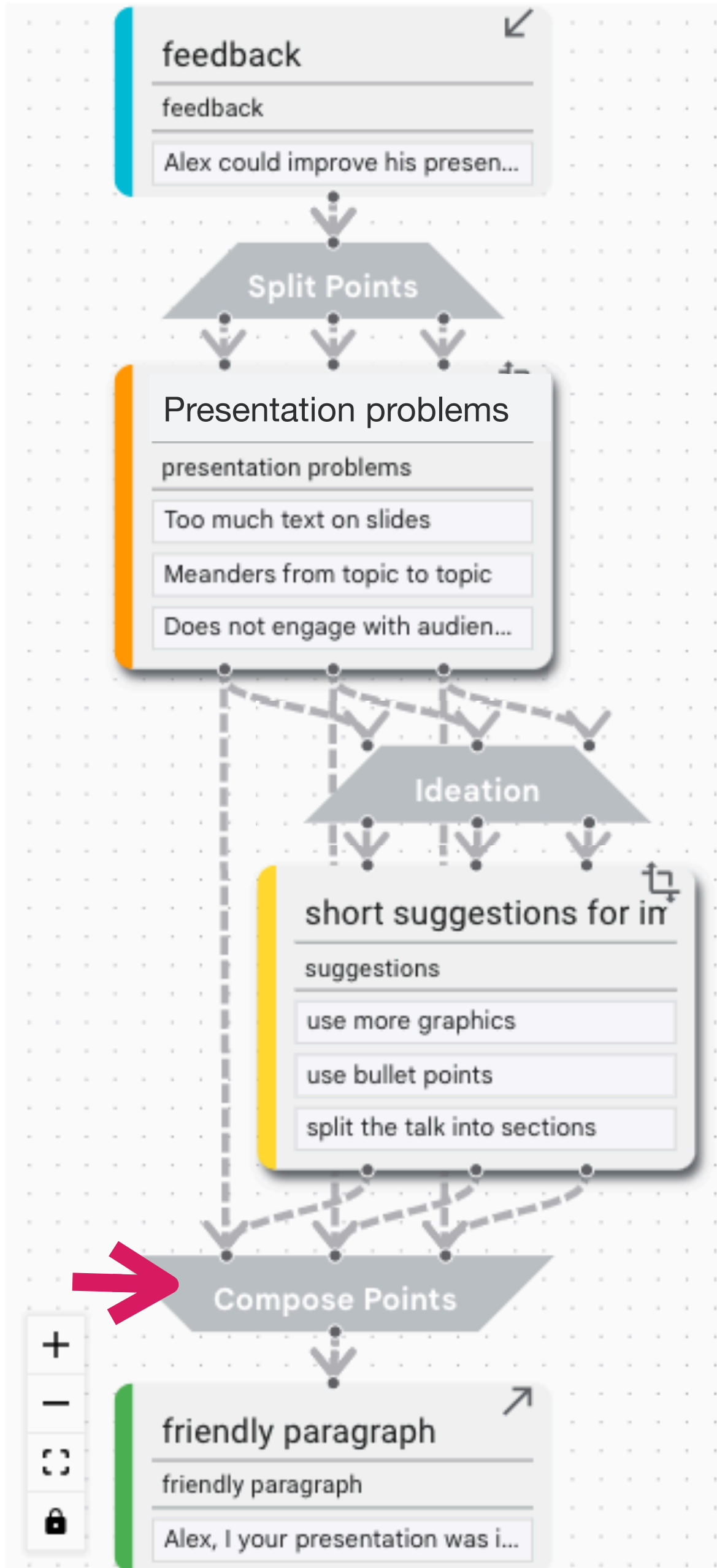
**Problem:** Meanders from topic to topic  
**Suggestions for improvements:**  
1) Split the talk into sections  
2) Read outlines and checkout each section as you go.



# A three-step Chain for peer review rewriting

3

## Compose the *final paragraph*



[[Instruction block]]

Write one friendly paragraph that covers all the presentation problems and suggestions.

[[Append running blocks]]

[[Running blocks]]

[[Prepend instruction blocks]]

**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 improvement** :

- 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.

Write one friendly paragraph that covers all the presentation problems and suggestions.

**Presentation problems:**

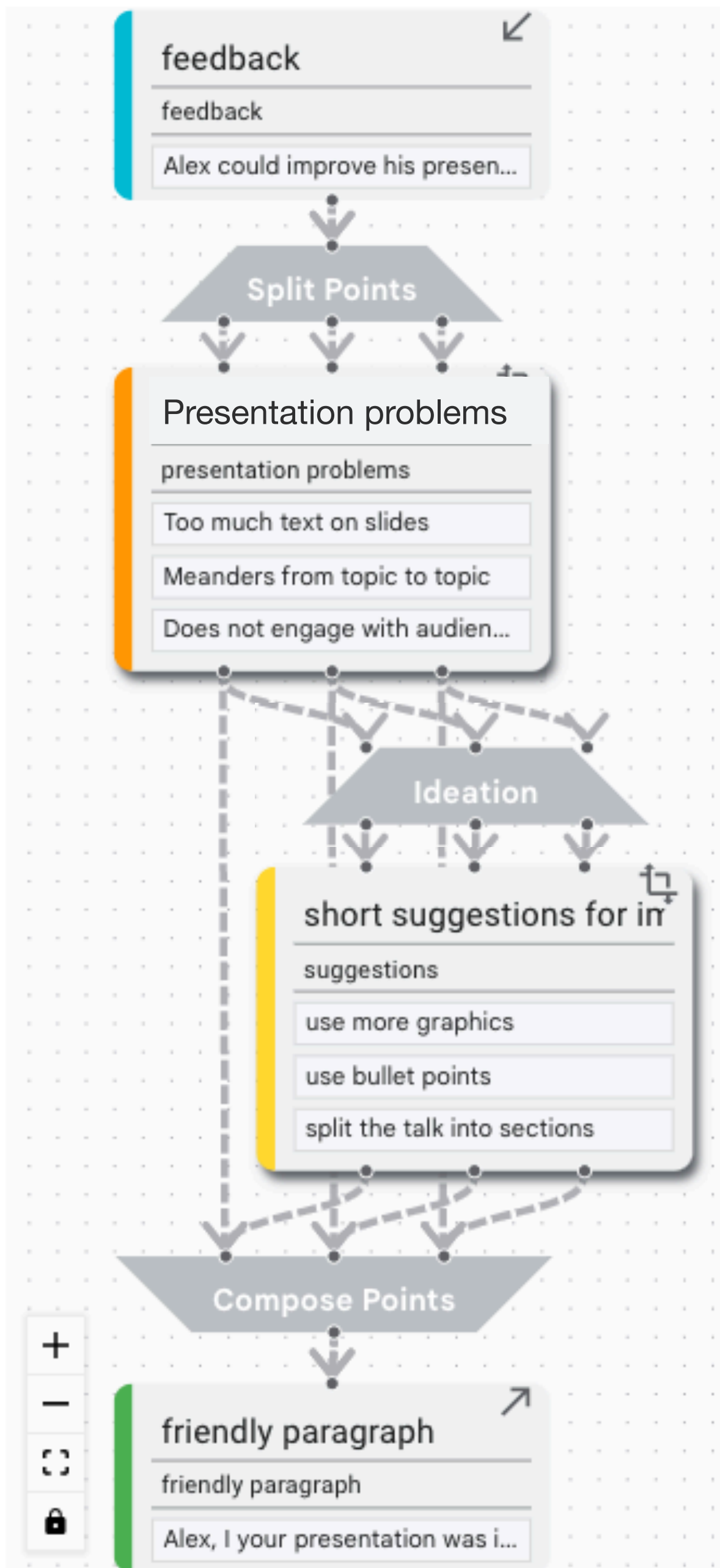
- 1) Too much text
- 2) Meanders from topic to topic

**Short suggestions for improvements:**

- 1) Use more graphics
- 2) Use bullet points
- 3) Split the talk into sections

**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.



[[Instruction block]]

Write one friendly paragraph that covers all the presentation problems and suggestions.  
 [[Append running blocks]]

[[Running blocks]]

[[Prepend instruction blocks]]

**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 improvement** :

- 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.

Less is more!

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



# Chaining improves human-AI interactions

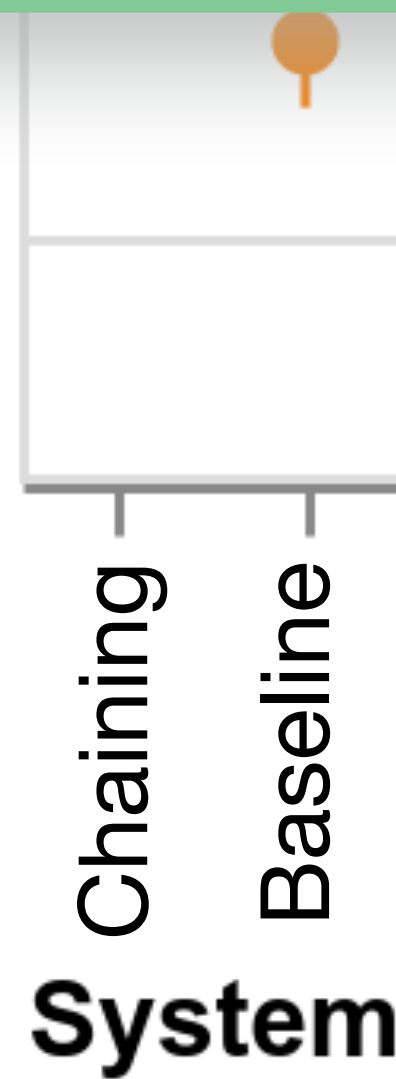
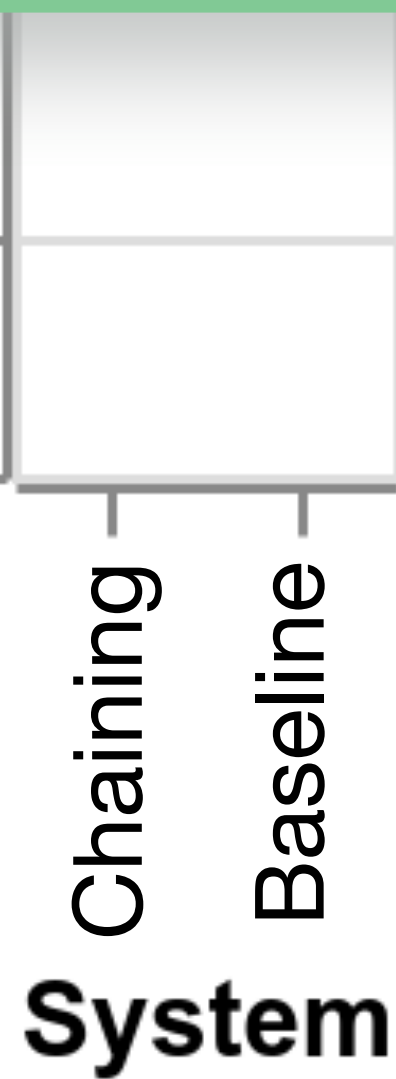
20

**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."*

Likert Scale

2  
0



## Rating Metrics

Transparent

Controllable

Collaborative

**System**  
● Chaining  
● Baseline

**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.

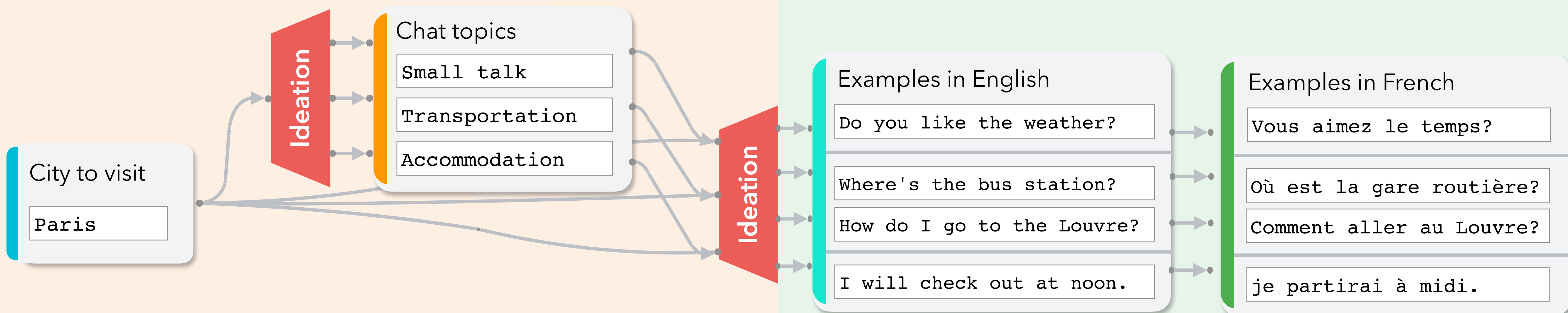
Exposure bias, overfit to their own generations.  
e.g., produce redundant content.

## Operations for LLM Chaining

Nudge users towards scoped inputs/outputs.

## Gather additional information from LLMs

e.g. Ideation



# 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.

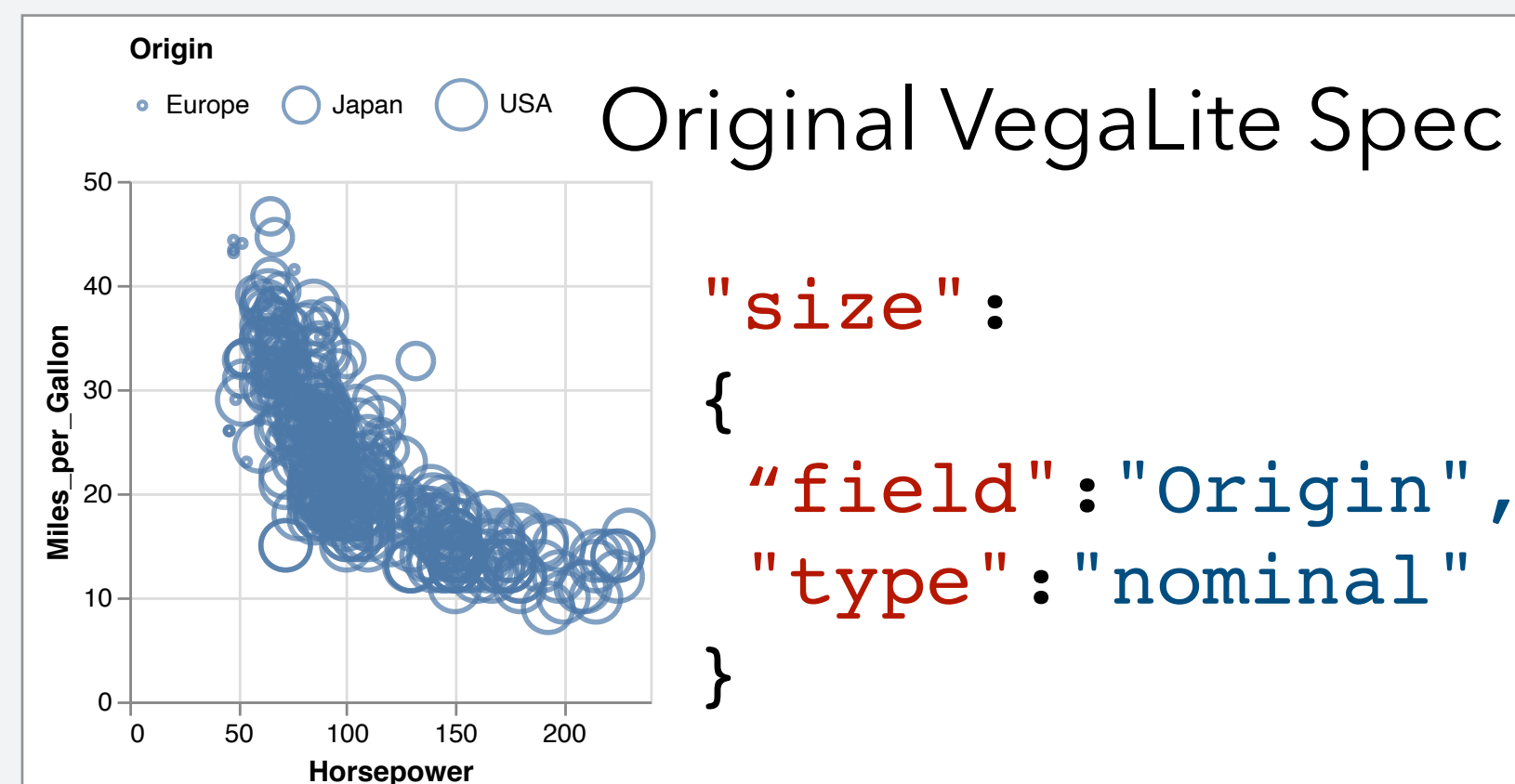
## Operations for LLM Chaining

Nudge users towards scoped inputs/outputs.

Gather additional information from LLMs

## Re-organize the input

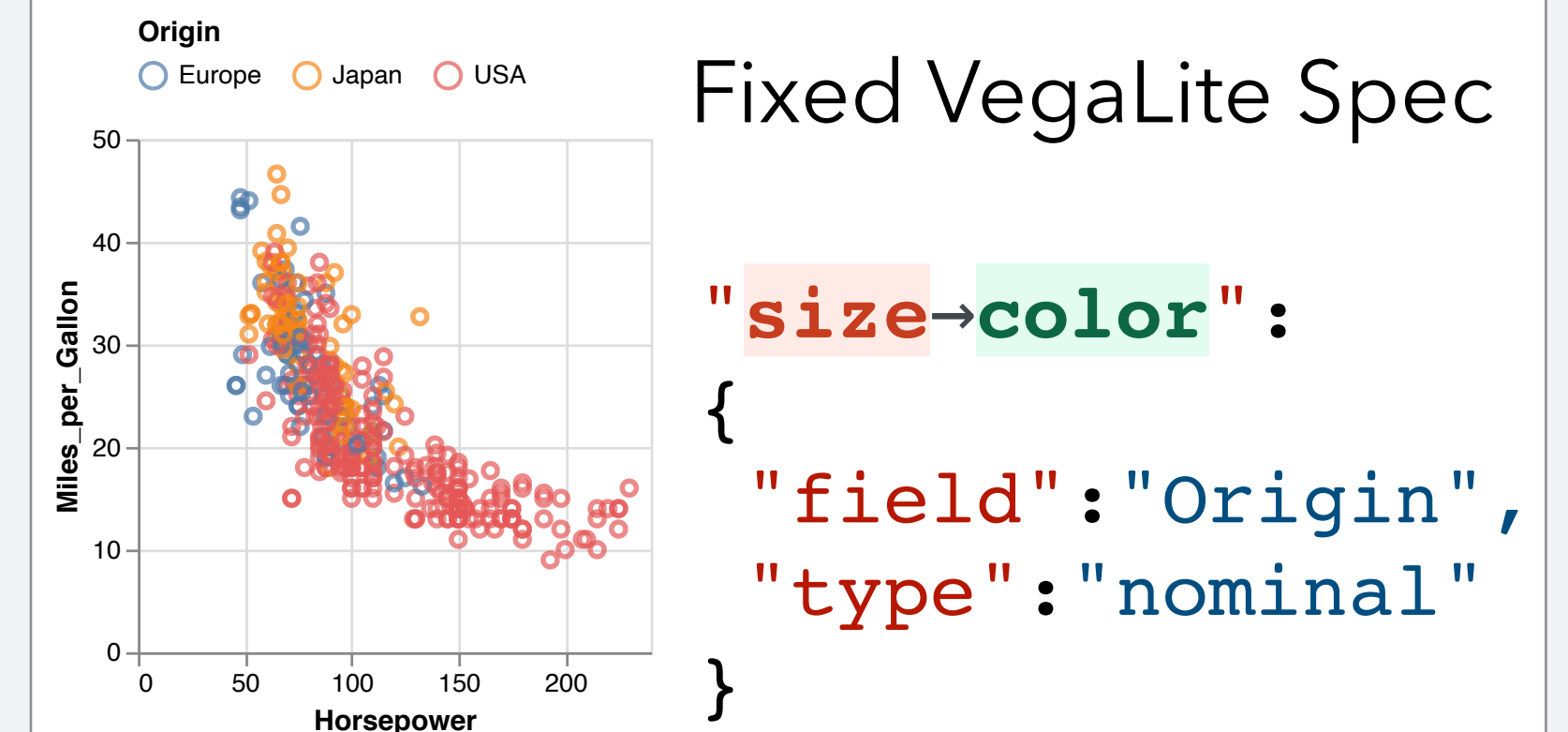
e.g. Rewriting



## Viz. description

```
A) [encoding:  
  size] has [data  
  type: nominal]  
B) ...
```

Rewriting



Rewriting

# 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

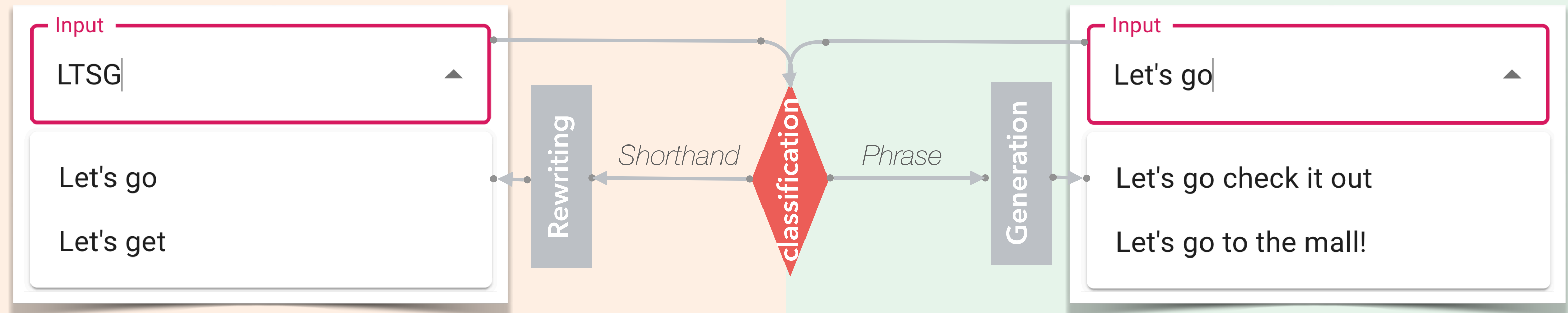
Nudge users towards scoped inputs/outputs.

Gather additional information from LLMs

Re-organize the input

## Validate and categorize the input

e.g. Classification



# 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

Play music by the Beatles.

Music play

I will play the following:

1. Get back
2. Hey Jude
3. Love me Do!

Who're some Country artists?

Music information

Garth Brooks, George Strait, ...

Hey! what up?

Not music

I'm chillin', how can I help you?

① Input PromptChainer: Chaining Large Language Model Prompts through

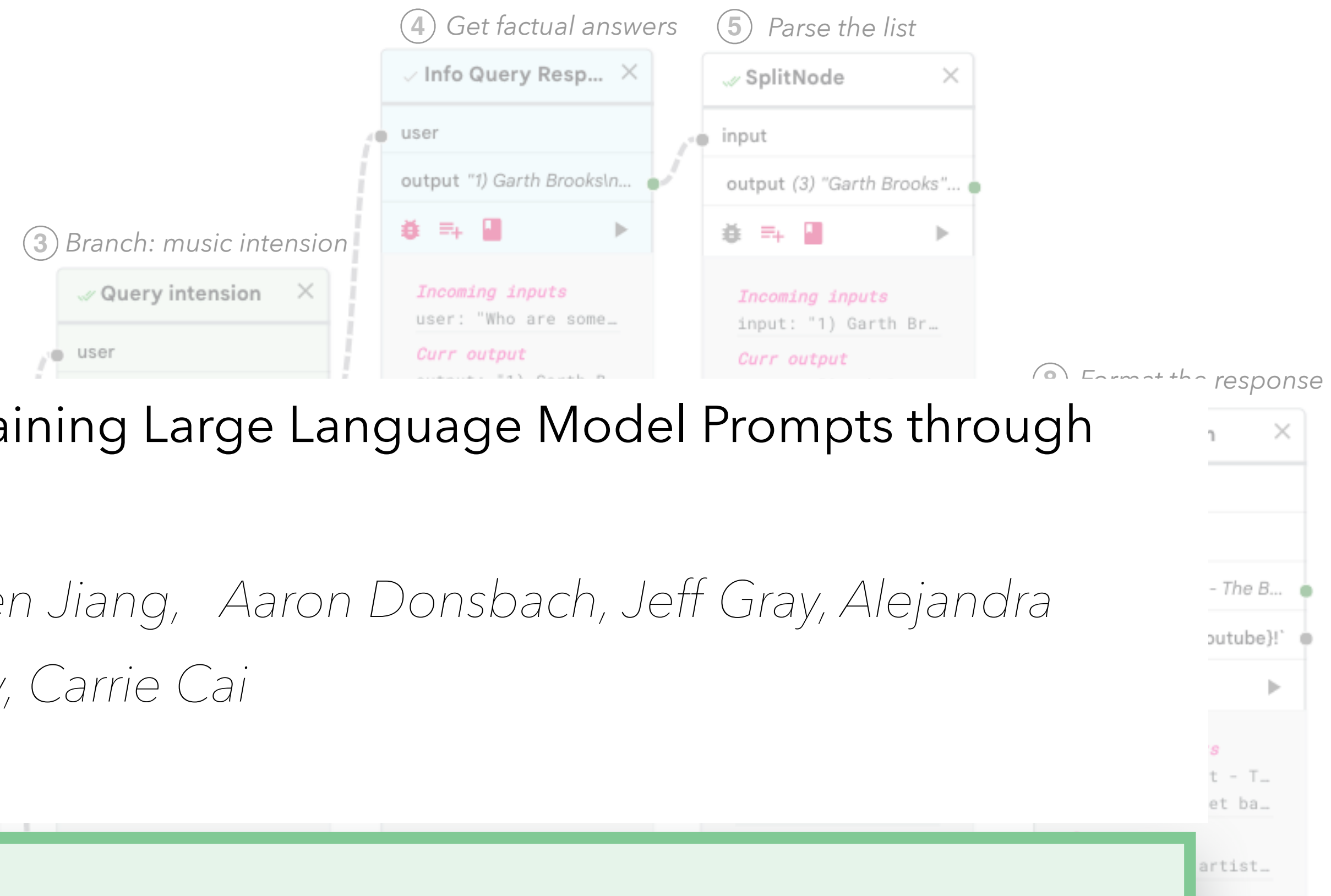
Visual Programming

*Tongshuang Wu, Ellen Jiang, Aaron Donsbach, Jeff Gray, Alejandra*

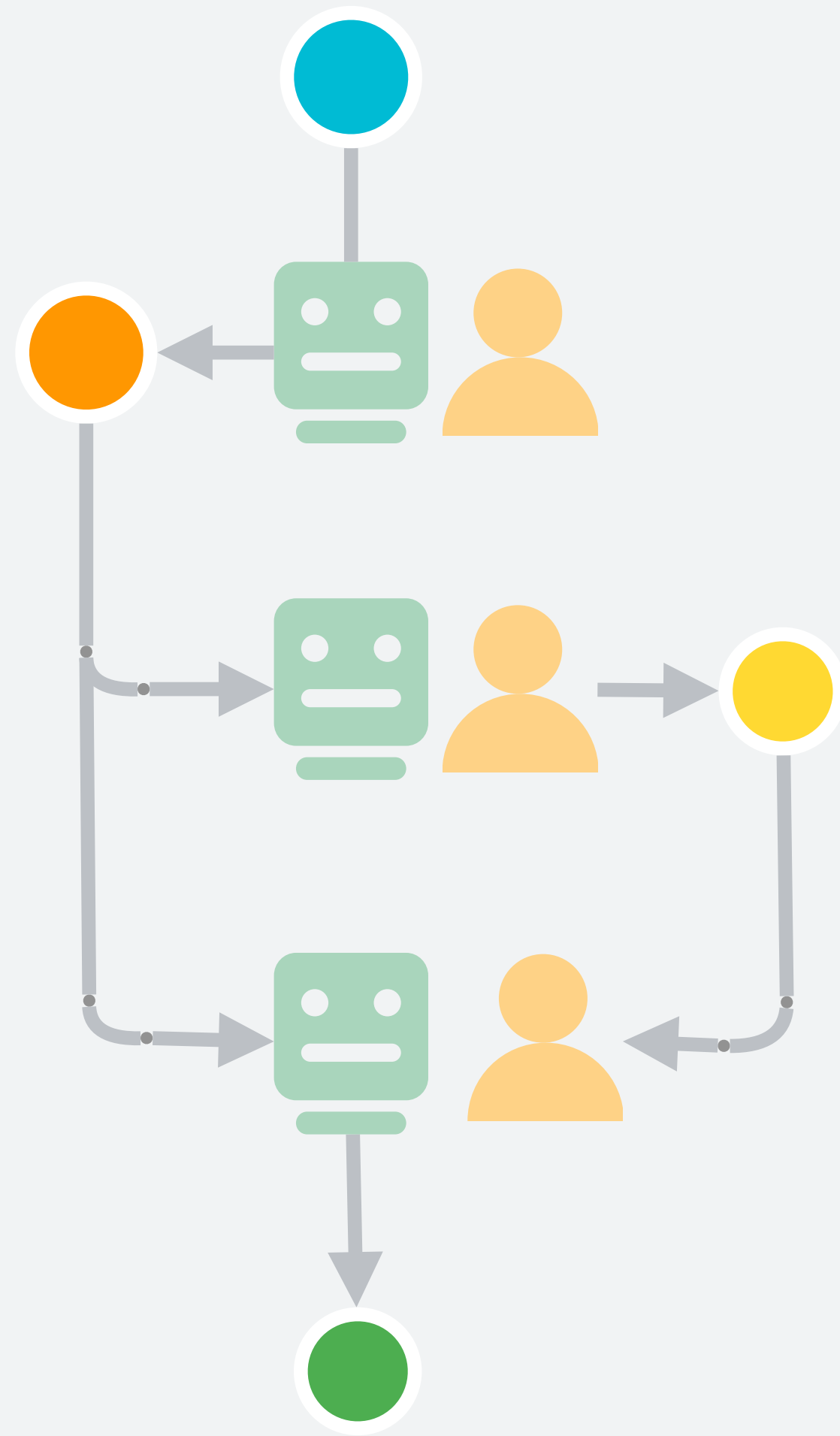
*Molina, Michael Terry, Carrie Cai*

*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

