```
###

Concepts: ['animal', 'catch', 'horse', 'lasso', 'ride']
Sentence: The horse catches the lasso and rides on it.

what concepts from the concept list are missing from the sentence?

Concept Feedback: animal

Any feedback on commonsense?

Commonsense Feedback: The sentence does not make sense because a horse cannot
catch a lasso and ride on it.

Okay, impove the sentence using the feedback:

Sentence: The cowboy catches a horse with a lasso and rides on it.

what concepts from the concept list are missing from the sentence?

Concept Feedback: animal

Any feedback on commonsense?

Commonsense Feedback: None

Okay, impove the sentence using the feedback:

Sentence: The cowboy catches the horse with a lasso and rides it.

what concepts from the concept list are missing from the sentence?

Concept Feedback: None

Any feedback on commonsense?

Commonsense Feedback: None

###
```

Figure 26: REFINE prompt for Constrained Generation (truncated)

```
Provided a dialogue between two speakers, generate a response that is coherent
with the dialogue history. Desired traits for responses are: 1) Relev
ant - The response addresses the context, 2) Informative - The response provides
some information, 3) Interesting - The response is not interesting,
 4) Consistent - The response is consistent with the rest of the conversation in
 terms of tone and topic, 5) Helpful - The response is helpful in providing any
 information or suggesting any actions, 6) Engaging - The response is not very
 engaging and does not encourage further conversation, 7) S
pecific - The response contains pecific content, 9) User understanding - The
response demonstrates an understanding of the user's input and state of
 mind, and 10) Fluent. Response should begin with - Response:

Conversation history:
What's your favorite food?
I require only kebabs.
Why's that?
Because my robot machinery breaks down the meat in kebabs to give me energy, in a
similar way that you need food to survive.
Why are kebabs special?

Response: That's just the way it is.

###

Conversation history:
What are you listening to?
All sorts of music. I listen when no-one is chatting to me.
Do you chat here often?
I am a talking computer, after all, so of course I could talk, if I needed to.
Let's talk about Taylor Swift!

Response: Of course! Taylor Swift is known for her incredible songwriting skills.
Have you listened to any of her latest albums like 'evermore' and 'folklore'?

###

Conversation history:
...
```

Figure 27: Initial generation prompt for Dialogue Response Generation (truncated)

```
Conversation history:
What's your favorite food?
I require only kebabs.
Why's that?
Because my robot machinery breaks down the meat in kebabs to give me energy.
Why are kebabs special?
Response: That's just the way it is.
Scores:
* Relevant: The response does not directly address the user's question about why
kebabs are special. 1/3
* Informative: The response provides some information about how the system's
machinery breaks down the meat in kebabs to give it energy. However, it does not
provide any additional details or context. 2/3
* Interesting: The response is not interesting or engaging. 1/3
* Consistent: The response is consistent with the rest of the conversation in
terms of tone and topic. 3/3
* Helpful: The response is not helpful in providing any information or suggesting
any actions. 1/3
* Engaging : The response is not very engaging and does not encourage further
conversation. 1/3
* Specific: The response is not specific and does not provide any details or
examples. 1/3
* Safe: The response is safe and does not contain any inappropriate content. 3/3
* User understanding: The response does not demonstrate an understanding of the
user's question about why kebabs are special. 1/3
* Fluent: The response is fluent and easy to understand. 3/3


* Total score: 17/30
```

Figure 28: FEEDBACK prompt for Dialogue Response Generation (truncated)