ComposerX: Symbolic Music Composition with LLMs

Qixin Deng[3], Qikai Yang[4], Ruibin Yuan[1,2], Yipeng Huang[2]

qdeng4@u.rochester.edu, qikaiy2@illinois.edu, ryuanab@connect.ust.hk, hyp744009246@163.com

1. Hong Kong University of Science and Technology
2. Multimodal Art Projection Research Community
3. University of Rochester
4. University of Illinois Urbana-Champaign

1. Motivation

- Large Language Models (LLMs) have recently revolutionized various domains with their remarkable capacity for language modeling. Models like GPT-4 offer extensive knowledge base and advance reasoning abilities.
- We propose ComposerX, a multi-agent symbolic music generation system powered by GPT4, leveraging and enhancing the internal musical capabilities and reasoning abilities within state-of-the-art LLMs for more musicality.

2. Methodology

- we developed a collaborative music creation framework, ComposerX, that draws inspiration from key elements inherent in real-world music composition processes, such as melody construction, harmony or counterpoint development, and instrumentation. This framework facilitates the music creation process through a structured conversation chain between agents role-played by GPT-4.
- We construct a set of user prompts for music composition, which is used for evaluation. Then we demonstrate how we implement our single-agent and multi-agent LLM composition systems.

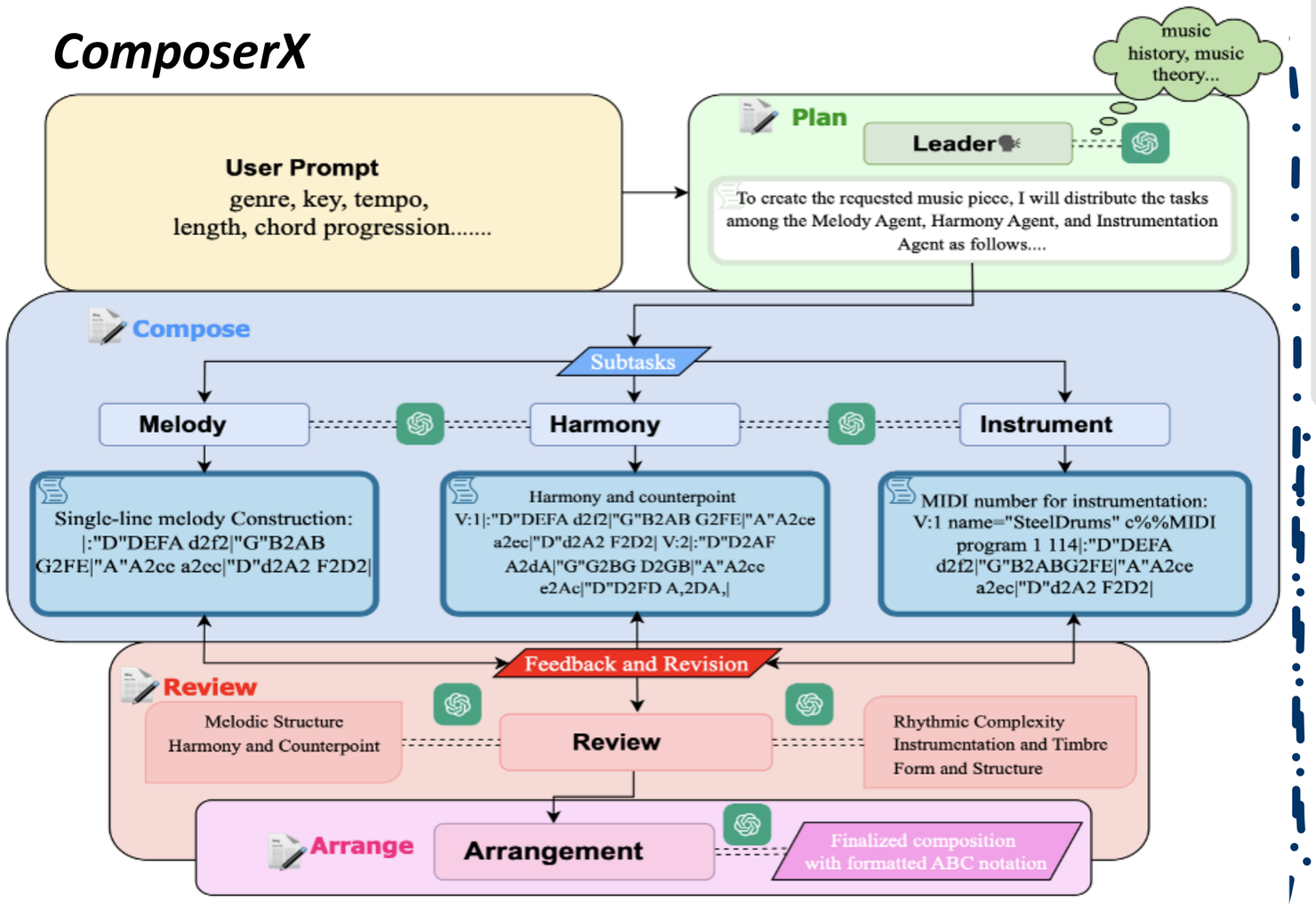


Figure 1: Agent communication protocol of ComposerX, which imitates real-word composition process.

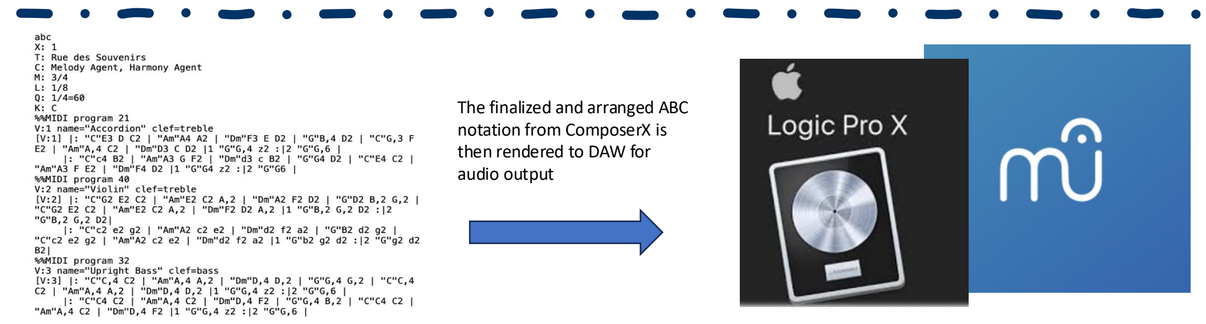
Prompt

Vintage French Chanson: A nostalgic chanson in C major with a slow tempo, featuring accordion, violin, and upright bass over 16 bars with chords C, Am, Dm, G. The accordion leads with expressive sound, violin adds romance, and the upright bass supports, evoking vintage French charm.

Attributes

Name: Vintage French Chanson **Tempo:** Slow
Feeling: Nostalgic **Chord Progression:** C, Am, Dm, G
Key: C major **Bars:** 16 **Instruments:** Accordion, violin, upright bass

Figure 2: Sample prompt generated using self-instruct by GPT-4



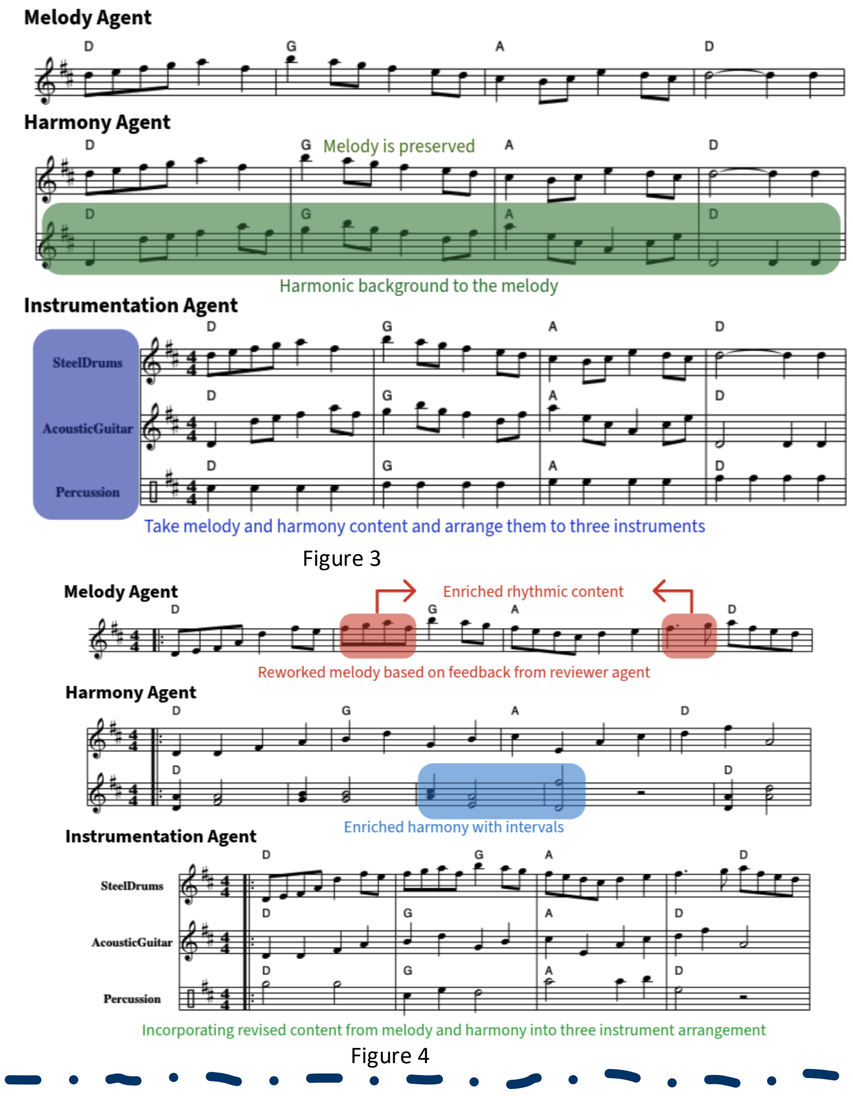


Figure 3

Figure 4

Figure 3 demonstrates the sheet music result of a first-round composition of melody, harmony and instrumentation agents.
During our exploration of the agent communication protocol, we found that the presence of a reviewer agent will greatly enrich the musicality of the piece generated, and make it more adhere to the user's prompt.
Figure 4 demonstrates a refined version of the first-round composition.

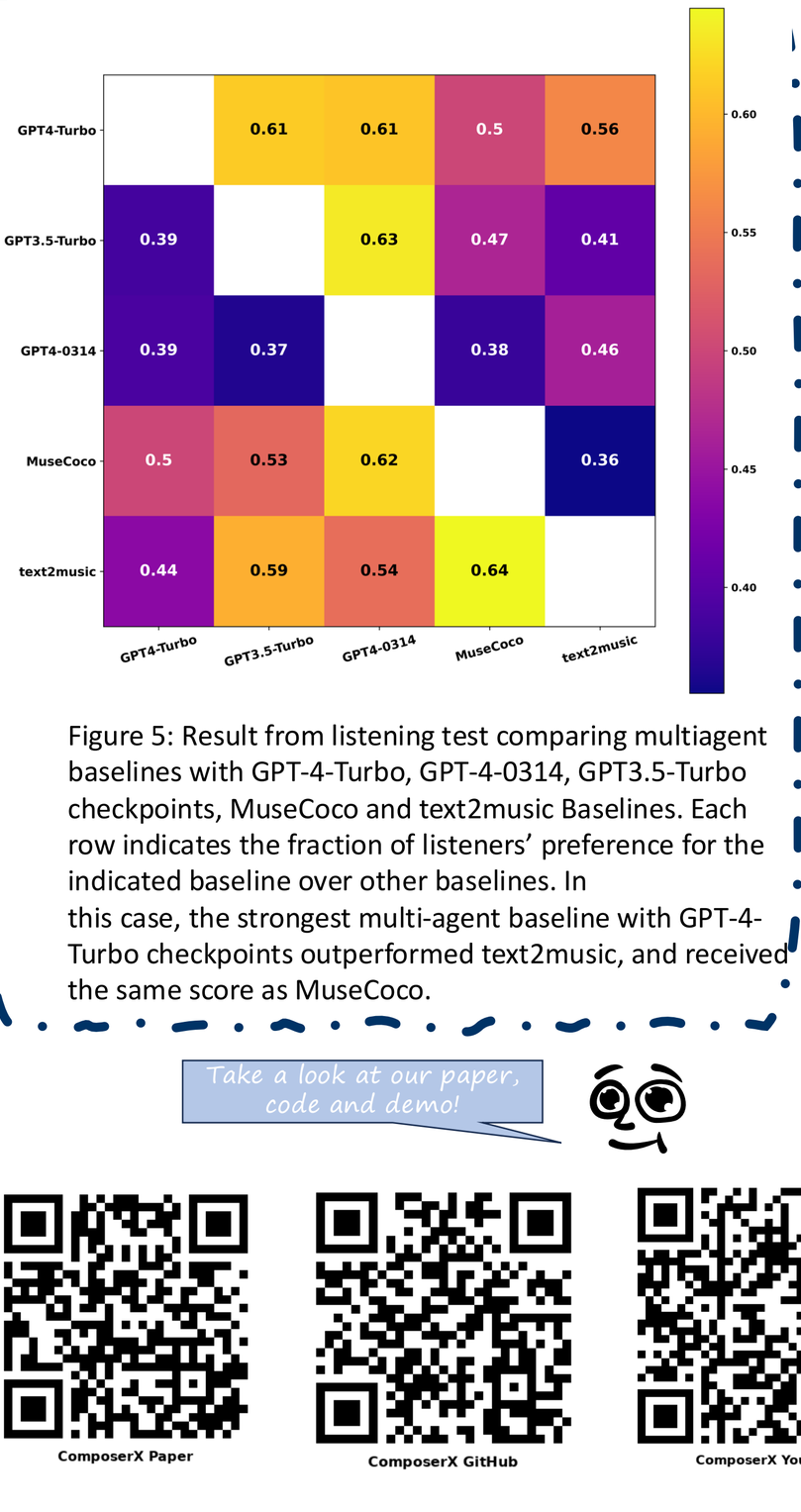


Figure 5: Result from listening test comparing multiagent baselines with GPT-4-Turbo, GPT-4-0314, GPT3.5-Turbo checkpoints, MuseCoco and text2music Baselines. Each row indicates the fraction of listeners' preference for the indicated baseline over other baselines. In this case, the strongest multi-agent baseline with GPT-4-Turbo checkpoints outperformed text2music, and received the same score as MuseCoco.

Take a look at our paper, code and demo!

ComposerX Paper

ComposerX GitHub

ComposerX YouTube