Bonus Assignment - Survey

Thank you for completing the bonus assignment! This survey will help us understand your experience, challenges, and suggestions for improving the framework. Your responses are invaluable for our research. Please answer the following questions thoughtfully.

**Section 1: Implementation Result**

1. **Please record the command you use to create, run, and evaluate the framework.**

*(For example, as provided in the README):*

- **Create**:

|  |
| --- |
| python create.py --config ./tz2634.json --output-dir ./args.output\_dir |

- **Run**:

|  |
| --- |
| python run.py --input-dir ./args.output\_dir |

- **Evaluation**:

|  |
| --- |
| python model\_api.py –-input-dir ./args.output\_dir  python eval.py \  --model\_api http://127.0.0.1:8000/eval/chat \  --config ./tz2634.json \  --documents\_dir ./args.output\_dir \  --output-dir ./args.output\_dir |

1. **Please copy and paste a sample conversation where you interacted with your created Agent.**

Bot: Hello JJ Lin fans! I'm your JJ Lin Fan Ambassador. How can I assist you today? Whether you're looking for the latest updates, exploring his music, or connecting with the fan community, I'm here to help!

You: What are some of the best songs of JJ Lin?

getAPIBotResponse Time: 6.182406902313232

Bot: JJ Lin has a vast repertoire of incredible songs that showcase his talent and versatility. Some of his best songs include:

1. "If You Don't Let Me Go" - A heartfelt ballad that highlights his emotional depth.

2. "Stories Untold" - Known for its captivating storytelling and melodic charm.

3. "Drifter" - A song that blends contemporary sounds with JJ's unique style.

4. "Twilight" - A beautiful track that showcases his powerful vocals.

5. "Practice Love" - A song with meaningful lyrics about love and relationships.

Here is a curated list of JJ Lin's albums and singles for you to explore. Enjoy diving into his music!

You: Tell me more about Drifter

getAPIBotResponse Time: 4.101596832275391

Bot: "Drifter" is one of JJ Lin's standout tracks, known for its contemporary sound and introspective lyrics. The song captures a sense of searching and self-discovery, with a blend of pop and R&B influences that highlight JJ's versatility as an artist. His emotive voice and the song's thoughtful composition resonate deeply with listeners, making it a favorite among fans.

If you're interested in exploring more about JJ Lin's music or have any other questions, feel free to ask!

You: Do I get a chance to listen to him live?

getAPIBotResponse Time: 9.11622405052185

Bot: Absolutely! JJ Lin is known for his incredible live performances and often connects with fans through concerts and world tours. His latest tour is ongoing, and he continues to captivate audiences around the globe. Keep an eye on his official website and social media channels for announcements about upcoming concert dates and locations. If there's a particular city or region you're interested in, I can help check if he has any scheduled performances there. Let me know how I can assist further!

You: I live in New York

getAPIBotResponse Time: 4.111940860748291

Bot: Great! JJ Lin often includes major cities like New York in his world tours. While I don't have the exact dates for his upcoming performances, I recommend checking his official website and social media platforms regularly for updates on concert dates and ticket availability. Attending one of his live shows is a fantastic way to experience his music and connect with fellow fans.

Additionally, feel free to join JJ Lin's fan community online to stay informed about any exclusive live streams or events that might be happening in your area. Let me know if there's anything else you'd like to know!

You: Sounds great! I would like to listen to some of his songs before going to a concert. That will make the live show more engaging to me. Any recommendations?

getAPIBotResponse Time: 13.128539085388184

Bot: Absolutely! Here are some song recommendations to get you familiar with JJ Lin's music before attending his live show:

1. \*\*"Little Big Us"\*\* - An inspiring anthem that celebrates unity and togetherness.

2. \*\*"She Says"\*\* - A touching ballad that showcases JJ's emotive vocals.

3. \*\*"River South"\*\* - A classic hit that beautifully blends traditional and modern sounds.

4. \*\*"Those Were the Days"\*\* - A nostalgic track with heartfelt lyrics.

5. \*\*"The Story of Us"\*\* - A song that captures the essence of storytelling through music.

These songs will give you a great mix of his style and range. Enjoy listening, and it will surely make the concert experience even more special!

Would you like to receive updates on JJ Lin's upcoming concerts or new releases?

You: Sure!

getAPIBotResponse Time: 4.276123046875

Bot: Great! To stay updated on JJ Lin's upcoming concerts and new releases, make sure to follow his official website and social media channels. This way, you'll be among the first to know about any exciting announcements or events.

Enjoy diving into his music, and have a fantastic time at his live show when he visits New York! If you have any other questions or need further recommendations, feel free to ask.

You: I'm good for now. Thank you!

getAPIBotResponse Time: 2.601824998855591

Bot: You're welcome! I'm glad I could help. If you ever have more questions or need assistance in the future, don't hesitate to reach out. Enjoy JJ Lin's music and have a wonderful time at the concert!

You: ^C

1. **Which of the following components did you use to create your Agent? *(Check all that apply.)***

* **✓** Interactive Panel to customize task planning
* **✓** Taskgraph Generation
* **✗** Worker Customization
* **✓** Evaluation

**Section 2: Feedback on Components**

1. **For the fields in the config file, Which field(s) in the config file were most useful for creating your Agent? Which field(s) did you find least important? Are there additional fields or features you would suggest adding to improve the Agent’s capabilities?**

I feel like the “user\_objective” and “rag\_docs” were the most useful for the creating the agent. The user objective is perfectly fulfilled according to the evaluation results, and I can clearly see that the responses generated by the agent containing information from the webpages I listed.

The field “domain” is not as important, maybe because the domain I’m defining is too narrow and does not provide too much additional information to the agent.

There aren’t really any specific fields I find complementary to the current layout; however, I suspect that the agent doesn’t stick well enough to the outline in the task graph. This is especially evidenced by the bad bot goal completion in the evaluation, where the goal is typically fulfilled at the rear parts of the outline. It could also be that the conversation did not progress to the points where the next tasks can be brought up. I believe that adding more flexibility or place a few emphases/more weights on certain important tasks can address this problem, so that the agent can complete the most important goals in some cases.

1. **Were the instructions for constructing the config file clear? Please specify areas where improvements are needed.**
   1. The pages contain a few grammar errors, which might confuse the readers. For example, in the sentence “Config files are the simple and recommended way to build your TaskGraph. A standard JSON document, …”, it is ambiguous what the “JSON document” refers to.
   2. In the README file, it says “Create a config file, similar to the project/AgentOrg/agentorg/orchestrator/examples/customer\_service\_config.json”, which specifies an outdated file directory. This line needs an update.
   3. For “task\_doc” and “rag\_docs”, the field “num” is not clearly defined. More details or examples could be provided. How could there be more than one website when I only provide one url? Does the program navigate deeper into the embedded urls on the given webpage?
   4. It might be helpful to connect the creation of a config file to the overall workflow of agent creation. A flowchart could work.
2. **How well did the generated task planning (without modification) cover the tasks you expected the Agent to handle?**

* All
* **Most**
* some
* Few
* None

1. **How useful were the steps generated in the task planning (without modification)? Rate their usefulness on a scale from 1 to 5.** *(1: Not Useful at All, 5: Very Useful)*

* 1
* 2
* **3**
* 4
* 5

Some of them are a bit redundant or too complicated.

1. **If you used the Interactive Panel for customizing task planning, how useful was it? Rate its usefulness on a scale from 1 to 5.** *(1: Not Useful at All, 5: Very Useful)*

* 1
* 2
* 3
* **4**
* 5

But I failed to add in more *tasks*, I could only add in more *steps*.

1. ***Follow up to question 8:* Based on your experience in customizing task planning: How much of the original task planning did you modify?**

* 80%+
* 60%+
* 40%+
* 20%+
* **<20%**

The program successfully generated some ideas that I didn’t think of, but they are very useful. However, some steps might be too complex and trimming is needed.

1. **If you designed your own worker, what are they and how were they used?**

I did not design my own worker.

1. **Follow up to question 10: How difficult did you find customizing the Worker framework? Rate the difficulty on a scale from 1 to 5.** *(1: Very Easy, 5: Very Difficult)*

* 1
* 2
* 3
* 4
* 5

1. **How useful was the user simulator for evaluating your Agent? Rate its usefulness on a scale from 1 to 5.**  *(1: Not Useful at All, 5: Very Useful)*

* 1
* 2
* 3
* **4**
* 5

1. **What issues did you identify in your Agent during evaluation, and how did you address them?**

The bot goal completion was bad. I think simplifying or clarifying the bot goal might help. But in the real world, there might be multiple and compound bot goals, and the current program is not strong enough at attaining them.

**Section 3: Feedback on Framework**

1. **How would you rate the overall difficulty of using the Agent First Organization framework?** *(1: Very Easy, 5: Very Difficult)*

* 1
* **2**
* 3
* 4
* 5

1. **What improvements would you suggest for the framework?**

It might help to add in more flexibility to the task graph instead of strictly following it. It might also help to evaluate the importance of each task and each step to allow the agent to skip a few when appropriate. This mechanism could be modeled with probability, just an idea.

1. **Additional comments or suggestions?**

Very interesting tool! It makes customized agents scalable.