Memory Matters: Interaction Experience with a Quizzing Bot

Arnob Chowdhury, Esha Dutta, Mitali Patil, Smruti Kshirsagar
Delft University of Technology

ABSTRACT

In this paper, we introduce Biblo, a novel mathematics quizzing agent. Biblo has been developed using Furhat remote API in Python. It also contains facial and emotional recognition modules which have been developed using OpenCV and Tensorflow. Through this paper we aim to answer the research question: "To what extent does a conversational agent with memory enhance user interaction experience when compared to an agent without memory?". In Biblo's case, the agent with memory (model B) retains the number of correctly and incorrectly answered questions between sessions, the last emotion of the user from previous session, and the last six general talk conversations. This data is used in subsequent sessions. In the agent without memory (model A), the same conversation takes place without any reference to the previous quizzes or general talk. Thirty participants were asked to interact with one of the models over two sessions. After these sessions, they filled out a questionnaire which evaluated their user interaction experience.

1 INTRODUCTION

Social agents find applications in various domains such as healthcare, personal assistance, and education. In [9], a compelling case is made for incorporating intelligent tutoring systems to improve performance in mathematics. Thus, we developed Biblo, a mathematics quizzing agent designed to provide users with a coherent and contextually rich experience and instill a sense of achievement. We hypothesise that incorporating long-term memory in a quizzing bot will result in users gaining a better interaction experience.

Biblo's reasoning process is modeled according to the Belief-Desire-Intention framework. Biblo possesses a mathematics knowledge base containing questions and answers which serves as the belief component. It also maintains information about the user such as performance history and emotional state. Biblo's desires include engaging the user in quizzes based on their preferences and past interactions. It also desires to recognise users emotions and respond appropriately. Its intentions involve selecting quiz questions based on the users preference and previous incorrect questions, giving hints if needed, and evaluating responses.

In a conversational agent, memory cannot be a straightforward database and requires more sophisticated storage solutions due to complexities. According to [15], conversations can be inherently context-sensitive, and demand a memory system that not only stores dialogue data but also preserves the context in which it was utilised. The fluidity of dialogues and changing relevance of information over time pose challenges for a simplistic database-like memory.

[20] states that conversations are rife with ambiguity and often extend beyond text to include non-textual elements such as images, facial expressions, and bodily gestures. Memory does not include storing only conversation data but also understanding and reacting appropriately to emotional cues, which goes beyond simple data storage. Addressing these complexities, especially in a tutoring agent use case, requires a memory system capable of contextual understanding and supporting multi-modality.

Biblo can ask questions about algebra, geometry, percentages and general mathematics. Biblo also takes into account whether the user is in the mood for solving mathematics questions through facial cues and verbal confirmations. If the user would rather not be quizzed, Biblo switches to general talk unrelated to mathematics. The agent is equipped with facial recognition. Moreover, it detects the emotions of the user and incorporates it into conversations. The source of Biblo's speech is predefined statements and natural language generation models. Additionally, a knowledge base of 50 mathematics questions has been created manually.

The content of speech spoken by Biblo consists of either mathematics questions or general talk. The context of the conversation is characterized by understanding the users' intent of solving questions, their category preferences, their emotions, and the history of the conversations with that user. Common ground with the user is established for effective communication by understanding their intents and emotions, and also shared knowledge regarding the correctness of their answers to the mathematics questions asked by Biblo. In this case, the knowledge base serves as semantic memory while user-specific conversational data serves as the episodic memory.

We have created two models – model A is a stateless conversational agent that relies on the current input to generate responses and model B contains the short-term and long-term memory architecture described above, and the research question we aim to answer is: "To what extent does a conversational agent with memory enhance the user interaction experience when compared to an agent without memory?"

More specifically, model B retains the number of correctly and incorrectly answered questions between sessions along with the last six general talk conversations. This data is used in subsequent sessions to revise the incorrect answers. In model A, the same conversation takes place without any reference to the quizzes or general talks in the previous sessions.

The remainder of this paper is structured in the following way: The Background section establishes context for the development of the conversational agent for mathematics quizzes. The Methodology section outlines the architectural design employed. The Results section presents the outcomes of the experiments conducted.

2 BACKGROUND

Long-term memory in conversational agents involves retaining information gathered from interactions over an extended period. In contrast, short-term memory, as defined by [11], is a temporary storage for information relevant to the current situation.

Significance of long-term memory lies in enabling conversational agents to build lasting relationships with users, integrating episodic

data over time. [14] highlights that recalling past interactions reflects caring and interest from the agents, fostering personalized interactions and sustained interest from users. The following paragraphs discuss several papers that explore the implementation of episodic memory in conversational agents.
Embodied Conversational Agents (ECAs) face the challenge of generating continuous, socially appropriate content. [1] suggests that to enhance user engagement, the conversational agent should actively communicate its intent to revisit specific topics or past interactions, thus fostering a collaborative interaction with the user.

Recall of past information contributes to making a more believable conversational agent. Notably, the addition of a perception component contributes emotional intelligence to the agent, enabling it to offer supportive feedback. According to [18], the ability of Interactive Virtual Agents (IVAs) to recall personal information is crucial, demonstrating care and interest by adjusting behavior throughout interactive sessions. It proposes to customise the chattiness of the IVA according to user profile and preference. There are correlations between the extent of recall/forgetfulness and the user's trust. Moreover, it was shown in [16] that older children had a positive experience and preferred talking to robots when the latter referred to prior shared events when communicating.

[11] proposes a conversational memory architecture for ECAs, extracting and storing relevant snippets to enhance intelligence. Utilizing APIs for named entity recognition, the architecture stores data in episodic memory, allowing agents to recall relevant information during conversations. Testing shows that ECAs with episodic memory perform more effectively. They show that an ECA with memory could answer questions more effectively than the same ECA without memory and that having episodic memory is an important step in the pursuit of making an efficient ECA.

In the study in [14], the authors focus on the importance of remembering past interactions for establishing long-term relationships with virtual characters. They create a system using Episodic Memory, Hierarchical Task Network Planning, and Belief-Desire-Intention architecture. Evaluation indicates that recalling past interactions sustains user interest over time.

[19] proposes a framework for episodic long-term memory in service robots. The model involves carefully consolidating short-term memories into long-term ones stored in a PostgreSQL database. Evaluation shows that users feel closeness and empathy with the robot, leading to a satisfactory experience.
These studies emphasize the importance of a contextual memory component in effective conversational agents. Long-term memory contributes to sustained user engagement and enhanced user experience.

3 METHODOLOGY

3.1 Architecture

Architecture for Biblo model A and model B are illustrated in figure 4 and 5 respectively.
Biblo includes speech and image as multimodal inputs. The facial recognition module has been created using the LBPHFaceRecognizer class in the OpenCV library in Python.[4] Images of a user's face are detected and stored in a folder. Their names and user IDs are stored in a local SQL table implemented using SQLite3 in Python.[5] The conversational agent queries this table to retrieve current user data. If a user is conversing with Biblo for the first time, their name and image data are added to the respective locations.

A pre-trained tensorflow[6] module for emotion recognition is executed after the facial recognition module successfully recognizes a returning user or adds a new user to the database. Throughout the session with a user, this model categorises the users emotions as one of the following: angry, disgusted, fearful, happy, neutral, sad, or surprised. The classified emotion is updated once every ten seconds in the aforementioned SQL table for the current user interacting with Biblo. This emotion data is queried periodically from the table during the session and sent in text format along with user's speech to an OpenAI Large Language Model. Accordingly, the model returns an empathising response accompanied with a change in facial expression.

The schema of the SQL table can be referred to in figure 1. The process is explained in figure 2.

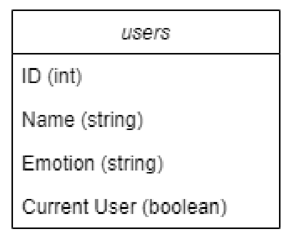

Figure 1: Schema of Users Table

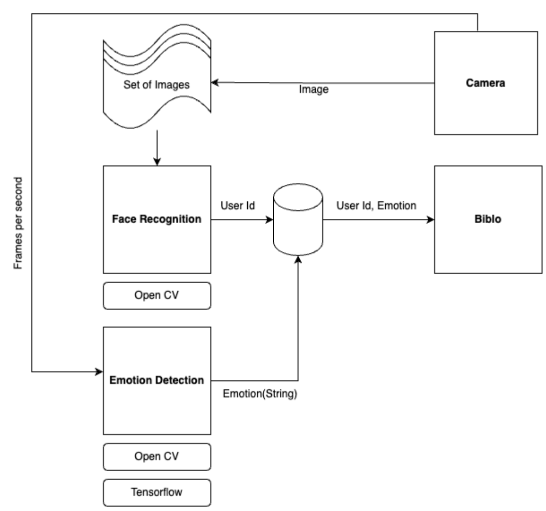

Figure 2: Facial Recognition and Emotion Detection Flow

Biblo asks questions by enunciating them vocally. The same questions and options are also printed on a user interface. This UI has been created using Tkinter library in Python.[7] This enables the user to reference the questions and options while solving the problem. The user can answer the question by selecting the correct answer on the UI and pressing the "Submit" button. There is also a "Hint" button if the user needs help with the question.

Facial recognition, emotion detection, and UI are implemented in both model A and model B.

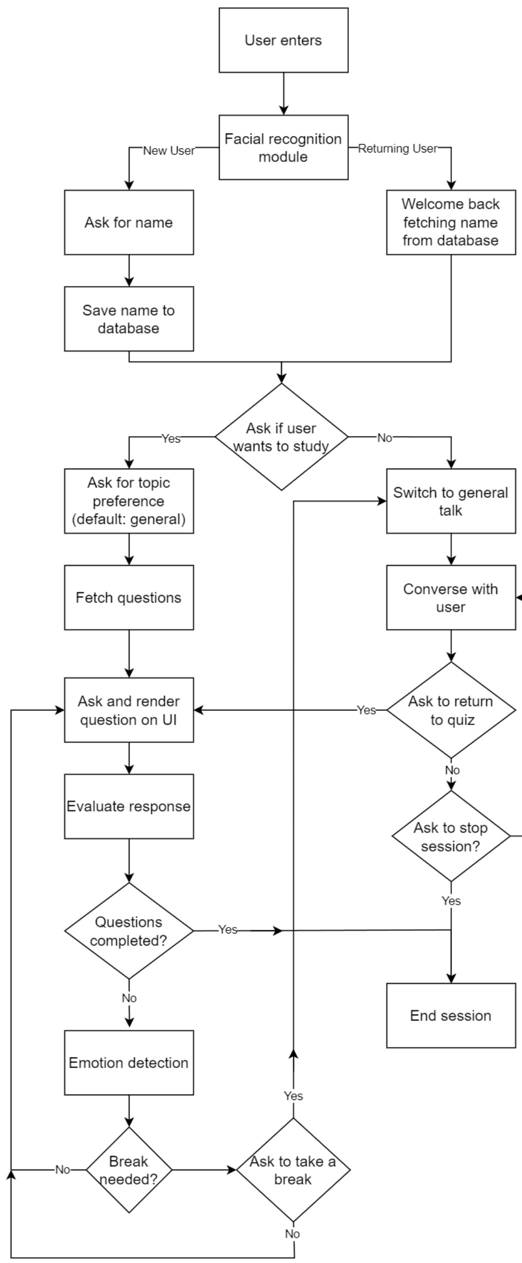

Figure 3: Session Flow Diagram

Biblo model B has three types of memory: short-term, long-term, and a knowledge base containing mathematics questions. The memory and knowledge base use a vector-based database provided by Pinecone[17], enabling the bot to perform a smart search based on semantics.

For short-term memory, during general talk, the bot searches for similarities in the sentences user provides with recent conversations. If the current conversation matches previous ones semantically, the bot fetches information based on both recency and similarity. Forgetting is simulated by utilizing only the last six general talk dialogues. If there's no semantic match, the bot relies on recency alone. This approach helps maintain the context of the conversation. For example, if a user mentions "it," the bot understands what "it" refers to and responds accordingly.

After a session ends, Biblo summarizes and stores user's quiz information, including details like correct and incorrect answers, general conversation topics, and the user's last emotion. This information is stored in long-term memory. General talk data is stored for a duration of one week, after which it is deleted.

For model A, which does not contain memory, the session data is not preserved. Questions are fetched from a JSON file that serves as the knowledge base. Only the name of the user is present in the SQLite database at the end of the session; the quiz data and general talk dialogues are not persistent in memory.

A Python application drives the dialogue throughout a session. The dialogue management technique, as illustrated in figure 3, is a hybrid approach. It is majorly FSM-based with some Natural Language Generation from OpenAI's LLM gpt3.5-turbo[8] accessed via REST APIs. Conversation is driven by the agent. Furhat uses Google Cloud Speech To Text[3] and Amazon Polly Text to Speech[2]. Text sent to Furhat API is synthesised as speech, and synchronised lip movements are added automatically. Furhat remote API is called from a Python application. This application processes the input given by user and generates appropriate responses. Biblo's input and output are in the form of speech through Furhat along with text on Tkinter UI.

After user is detected through facial recognition, the Python application interacts with them through Furhat. Users are asked to give their name and topic of preference. Input is processed using Google's language model – flan-t5-base[10]. Passing an appropriate prompt to Flan helps extract relevant information from the user's speech. The application operates in two contexts, namely quiz and general. Natural Language Understanding for quiz context is performed by this model. The Python application queries the Knowledge Base and fetches questions. In model B, if user has incorrectly answered questions from prior sessions for the same topic, these will also be fetched from long-term memory along with previously unanswered questions. This is not done in model A. At the end of the session, Biblo gives a summary of correct and wrong answers given by the user. After each question, depending on the correctness and current emotion of the user, Biblo sends a prompt to the OpenAI LLM to infer whether user wants to take a break, and asks the user accordingly. If the user agrees, the context switches to general. The user can converse with Biblo on any topic during this time before reverting to the quiz. Dialogue in general context is driven by OpenAI LLM. In model B, the past dialogues and quiz data from the same session stored in short-term memory are also sent to the LLM to give a more appropriate and nuanced response.

Some challenges were encountered during the creation of Biblo. There were some technical issues with Furhat remote API wherein it would get logged out during the session. The Furhat SDK was updated and installed again to solve this issue. Moreover, at times, Furhat misinterprets speech input when the user's enunciation is unclear. This was tackled by error handling and asking confirmation questions to ensure that correct input was recognized.

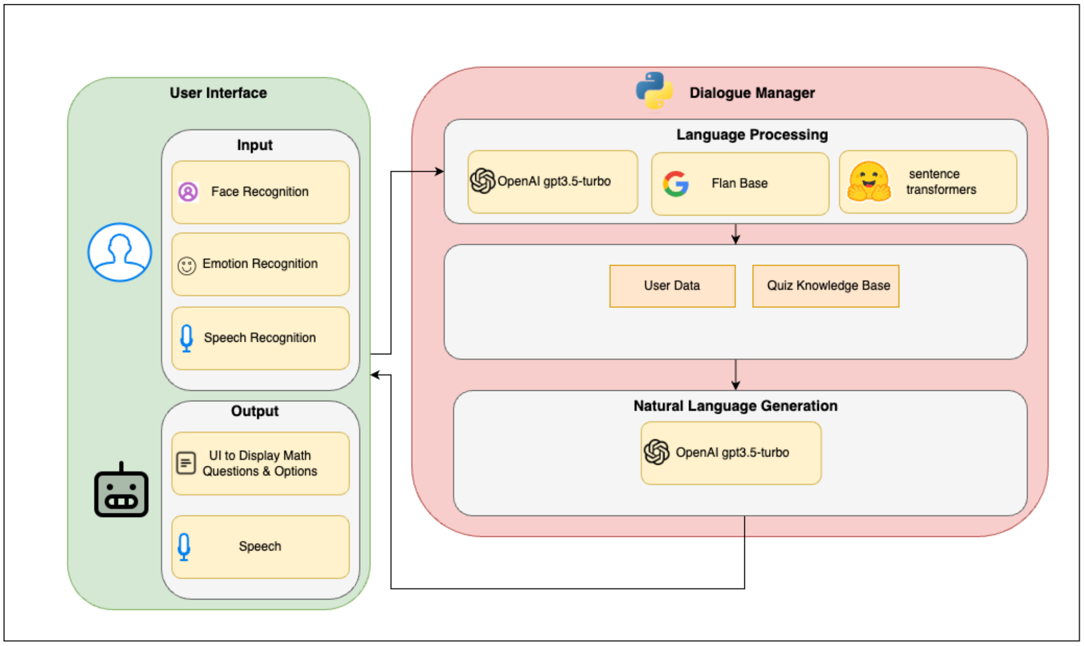

Figure 4: Biblo Model A Architecture

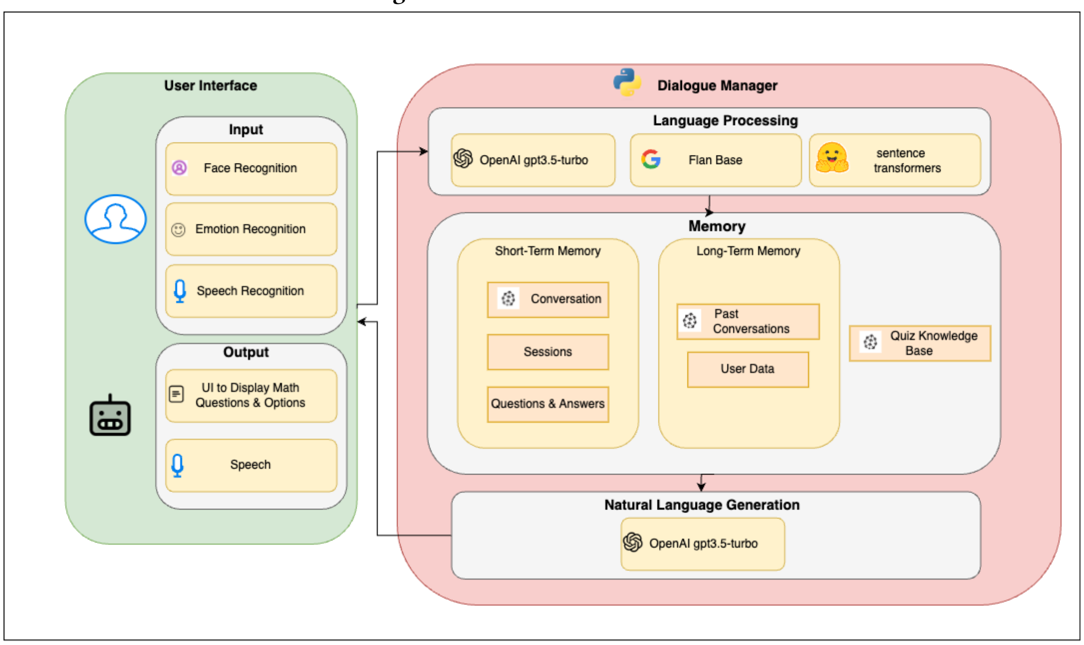

Figure 5: Biblo Model B Architecture

3.2 Experiment

The study was conducted between subjects. Fifteen participants interacted with model A and the other fifteen interacted with Model B. As a pre-test, we asked participants if they have knowledge of basic arithmetic operations. This ensured their preparedness to answer the questions. If the participant could not answer basic arithmetic questions they were not considered for the experiment and their respective pre-test data was discarded. When a participant arrived, they were asked to sign the consent form and began the experiment on an ASUS TUF A17 laptop. They were provided with a pen and paper to solve the mathematics questions. The face recognition model identified whether the user was new or returning. In case of a new user, Biblo asked the user's name. The model introduced itself and its purpose. Biblo initiated the conversation which the participant continued. Throughout this process, the emotion detection system was identifying the users emotions and saving it to a local database. The bot verbalized the question being displayed on the screen and the participant selected the answer to each question on the UI. There were two distinct sessions – the participant was recorded as a new user in the first session and as a returning user

Construct	U-value	P-Value	H_0 Rejected	Effect Size
Performance	730.5	0.004	Yes	0.27
User's Trust	815	0.043	Yes	0.18
User Engagement	733.5	0.007	Yes	0.26
User Agent Alliance	2722	0.000	Yes	0.30
Agent Coherence	1470.5	0.031	Yes	0.17
Agent Intentionality	1659	0.210	No	0.07

Table 1: Mann-Whitney U Test Results for Model A and B

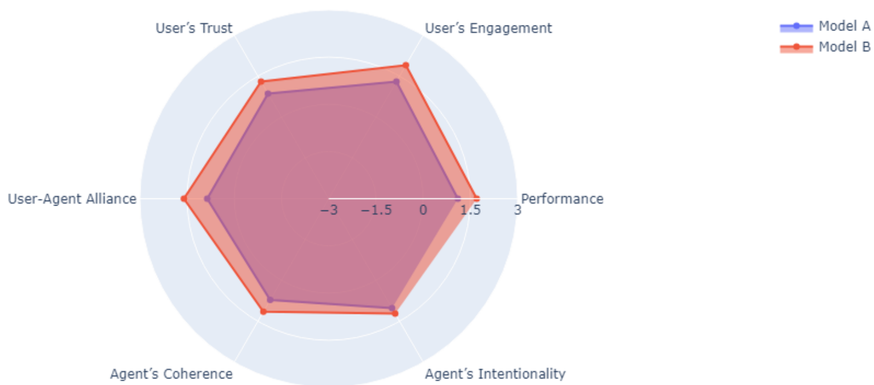

Figure 6: ASA Questionnaire Results

in the second session. Each session was 5-7 minutes long in which 3 mathematics questions were asked.
After these sessions, the participant data was discarded. Thereafter, the participants were asked to fill the questionnaire to evaluate their experience. This questionnaire serves as a common base to measure and describe the interaction experience with an artificial social agent [12]. It has 23 questions grouped into 6 categories. Both models have been assessed on Agent's Performance, User's Engagement, User's Trust of the Agent, User-Agent Alliance, Agent's Coherence and Agent's Intentionality.

4 RESULTS

We used a Likert scale to collect responses from participants. This scale facilitates the measurement of participants' opinions. The questions can be referred to in table 2. Each participant answered the questions on a scale from -3 to 3, where -3 means they strongly disagree, -1.5 means they disagree, 0 means they are neutral, 1.5 means they agree, and 3 means they strongly agree[12]. The measure was adjusted for questions with reverse scaling. The mean of the results of the questionnaire are visualized in 6.

We had planned to conduct an independent samples t-test. However, since the data is not normally distributed, a Mann-Whitney U Test was conducted[13]. None of the assumptions of the test are violated in our data. This test was left-tailed as model B is expected to display better user interaction experience. The type of model is the independent variable. The dependent variables are the constructs chosen from ASAQ. The hypotheses are as follows:
H_0 : the two models provide equal interaction experience
H_1 : interaction experience of model A < interaction experience of model B

The significance level, α, is chosen as 0.05 which is a commonly used significance level. The test was performed separately for the six constructs from the ASAQ.

In Agents Intentionality, since the p-value > α, the null hypothesis H_0 is not rejected. In all other constructs, the p-value < α and thus, the null hypothesis H_0 is rejected. The results of the test are summarized in 1.

The test reveals a difference between model A and model B for all constructs except for agent's intentionality.

5 DISCUSSION (EXAM)

6 CONCLUSION (EXAM)

REFERENCES

[1] 2015. *Lecture Notes in Computer Science* (Feb 2015), 78–87. https://doi.org/10.1007/978-3-319-15557-9
[2] 2024. AWS Polly TTS. https://aws.amazon.com/polly/. (2024). accessed on 21 Jan 2024.
[3] 2024. Google Cloud Speech-to-Text. https://cloud.google.com/speech-to-text/?hl=en. (2024). accessed on 21 Jan 2024.
[4] 2024. OpenCV documentation. https://docs.opencv.org/3.4/df/d25/classcv_1_1face_1_1LBPHFaceRecognizer.html. (2024). accessed on 21 Jan 2024.

[5] 2024. OpenCV documentation. https://docs.python.org/3/library/sqlite3.html. (2024). accessed on 21 Jan 2024.
[6] 2024. OpenCV documentation. https://www.tensorflow.org/api_docs/python/. (2024). accessed on 21 Jan 2024.
[7] 2024. Python Tkinter documentation. https://docs.python.org/3/library/tkinter.html. (2024). accessed on 21 Jan 2024.
[8] 2024. Text generation - openai API - platform.openai.com. https://platform.openai.com/docs/guides/text-generation. (2024). accessed on 21 Jan 2024.
[9] Royer James M. Arroyo, Ivon and Beverly P. Woolf. 2011. Using an Intelligent Tutor and Math Fluency Training to Improve Math Performance. In *International Journal of Artificial Intelligence in Education*.
[10] Hyung Won Chung, Le Hou, Shayne Longpre, Barret Zoph, Yi Tay, William Fedus, Yunxuan Li, Xuezhi Wang, Mostafa Dehghani, Siddhartha Brahma, Albert Webson, Shixiang Shane Gu, Zhuyun Dai, Mirac Suzgun, Xinyun Chen, Aakanksha Chowdhery, Alex Castro-Ros, Marie Pellat, Kevin Robinson, Dasha Valter, Sharan Narang, Gaurav Mishra, Adams Yu, Vincent Zhao, Yanping Huang, Andrew Dai, Hongkun Yu, Slav Petrov, Ed H. Chi, Jeff Dean, Jacob Devlin, Adam Roberts, Denny Zhou, Quoc V. Le, and Jason Wei. 2022. Scaling Instruction-Finetuned Language Models. (2022). arXiv:cs.LG/2210.11416
[11] Miguel Elvir, Avelino J. Gonzalez, Christopher Walls, and Bryan Wilder. 2017. Remembering a Conversation – A Conversational Memory Architecture for Embodied Conversational Agents. *Journal of Intelligent Systems* 26, 1 (2017), 1–21. https://doi.org/doi:10.1515/jisys-2015-0094
[12] Siska Fitrianie, Merijn Bruijnes, Fengxiang Li, Amal Abdulrahman, and Willem-Paul Brinkman. 2022. The artificial-social-agent questionnaire. *Proceedings of the 22nd ACM International Conference on Intelligent Virtual Agents* (2022). https://doi.org/10.1145/3514197.3549612
[13] Owen Ingram. 2023. Which statistical test should you use? (Aug 2023). https://www.researchprospect.com/which-statistical-test-you-should-use/
[14] Zerrin Kasap and Nadia Magnenat-Thalmann. 2011. Building long-term relationships with virtual and robotic characters: The role of remembering. *The Visual Computer* 28, 1 (2011), 87–97. https://doi.org/10.1007/s00371-011-0630-7
[15] Ludovic Le Bigot, Cléo Bangoura, Dominique Knutsen, and Sandrine Gil. 2021. When non-salient information becomes salient in conversational memory: Collaboration shapes the effects of emotion and self-production. *Quarterly Journal of Experimental Psychology* 75, 7 (2021), 1330–1342. https://doi.org/10.1177/17470218211055005
[16] Iolanda Leite, André Pereira, and Jill Fain Lehman. 2017. Persistent Memory in Repeated Child-Robot Conversations. In *Proceedings of the 2017 Conference on Interaction Design and Children (IDC '17)*. Association for Computing Machinery, New York, NY, USA, 238–247. https://doi.org/10.1145/3078072.3079728
[17] Pinecone. 2024. Pinecone Documentation. (2024). https://docs.pinecone.io/ Tue Jan 23, 2024.
[18] Deborah Richards and Karla Bransky. 2014. ForgetMeNot: What and how users expect intelligent virtual agents to recall and forget personal conversational content. *International Journal of Human-Computer Studies* 72, 5 (2014), 460–476. https://doi.org/10.1016/j.ijhcs.2014.01.005
[19] María-Loreto Sánchez, Mauricio Correa, Luz Martínez, and Javier Ruiz-del Solar. 2015. An Episodic Long-Term Memory for Robots: The Bender Case. In *RoboCup 2015: Robot World Cup XIX*, Luis Almeida, Jianmin Ji, Gerald Steinbauer, and Sean Luke (Eds.). Springer International Publishing, Cham, 264–275.
[20] W. Gill Woodall and Joseph P. Folger. 1985. Nonverbal cue context and episodic memory: On the availability and endurance of nonverbal behaviors as retrieval cues. *Communication Monographs* 52, 4 (1985), 319–333. https://doi.org/10.1080/03637758509376115

A APPENDIX

Question Bank

Algebra

(1) Solve the equation: $2x - 5 = 7$.
  (a) 2
  (b) 4
  (c) 6
  (d) 8
  (e) 10
(2) Simplify: $3(2x - 5) + x = 20$.
  (a) 2.5
  (b) 4.375
  (c) 5
  (d) 6
  (e) 7.5
(3) Factor: $x^2 + 2x - 4 - 8$.
  (a) $x + 2$
  (b) $x - 2$
  (c) $x^2 + 2$
  (d) $x - 2$
  (e) $x^2 - 2$
(4) Solve for x: $2(3x - 1) = 22 - x$.
  (a) 4
  (b) 5
  (c) 6
  (d) 7
  (e) 8
(5) Factorize: $x^2 - 5x + 6 = 0$.
  (a) 1
  (b) 2
  (c) 4
  (d) 5
  (e) 2 and 3
(6) Solve the quadratic equation: $2x^2 - 5x - 3 = 0$.
  (a) 1 and 2
  (b) -1 and 3
  (c) 3 and $-\frac{1}{2}$
  (d) $\frac{1}{2}$ and 2
  (e) -2 and 1
(7) Identify the vertex of the quadratic equation $y = 2x^2 - 8x + 6$.
  (a) $(1, 1)$
  (b) $(2, 2)$
  (c) $(-2, 6)$
  (d) $(4, -2)$
  (e) $(0, 4)$
(8) Simplify: $-2(x - 2) + 15$.
  (a) $-2x - 5$
  (b) $-2x + 5$
  (c) $-x + 5$
  (d) $x - 5$
  (e) $x + 5$
(9) Solve the quadratic equation: $3x^2 - 5x + 2 = 0$.
  (a) $\frac{1}{3}$ and 2
  (b) -1 and 3
  (c) 3 and $-\frac{1}{2}$
  (d) $\frac{1}{2}$ and 2
  (e) -2 and 1
(10) Factorize: $x^2 - 5x + 6$.
  (a) $x + 2$
  (b) $x - 2$
  (c) $x^2 + 2$
  (d) $(x - 2)(x - 3)$
  (e) $x^2 - 2$

Geometry

(1) Find the length of the hypotenuse in a right-angled triangle with legs of length 3 and 4.
  (a) 3
  (b) 5
  (c) 7
  (d) 9

(e) 12

(2) Find the area of a rectangle with length 6 units and width 8 units.
  (a) 12 square units
  (b) 24 square units
  (c) 48 square units
  (d) 64 square units
  (e) 96 square units

(3) Find the circumference of a circle with a radius of 6 units.
  (a) 12π units
  (b) 6π units
  (c) 9π units
  (d) 16π units
  (e) 24π units

(4) Find the volume of a cylinder with a radius of 3 units and height of 8 units.
  (a) 18π cubic units
  (b) 36π cubic units
  (c) 54π cubic units
  (d) 72π cubic units
  (e) 90π cubic units

(5) Find the surface area of a sphere with a radius of 3 units.
  (a) 9π square units
  (b) 18π square units
  (c) 27π square units
  (d) 32π square units
  (e) 36π square units

(6) Find the area of a triangle with a base of 5 units and a height of 4 units.
  (a) 5 square units
  (b) 10 square units
  (c) 15 square units
  (d) 20 square units
  (e) 25 square units

(7) Find the perimeter of a rectangle with length 12 units and width 5 units.
  (a) 26 units
  (b) 30 units
  (c) 34 units
  (d) 38 units
  (e) 42 units

(8) Find the area of a circle with a radius of 6 units.
  (a) 36π square units
  (b) 18π square units
  (c) 72π square units
  (d) 45π square units
  (e) 54π square units

(9) Find the volume of a cone with a radius of 3 units and height of 4 units.
  (a) 12π cubic units
  (b) 24π cubic units
  (c) 30π cubic units
  (d) 36π cubic units
  (e) 48π cubic units

(10) Find the diagonal length of a rectangle with length 5 units and width 12 units.
  (a) 8 units
  (b) 10 units
  (c) 11 units
  (d) 12 units
  (e) 13 units

Percentages

(1) Calculate the percentage increase when an amount increases from $50 to $70.
  (a) 20%
  (b) 40%
  (c) 50%
  (d) 30%
  (e) 60%

(2) Calculate the percentage decrease when an amount decreases from $120 to $80.
  (a) 25%
  (b) 30%
  (c) 33.33%
  (d) 40%
  (e) 50%

(3) What percentage is a score of 80 out of a maximum score of 200?
  (a) 20%
  (b) 30%
  (c) 35%
  (d) 40%
  (e) 45%

(4) Calculate the original value when an amount increases by 25
  (a) 100
  (b) 110
  (c) 115
  (d) 120
  (e) 130

(5) Calculate the final value when an amount increases by 20% from $150.
  (a) 160
  (b) 180
  (c) 190
  (d) 200
  (e) 220

(6) Calculate the original quantity when it decreases by 15% to reach 85.
  (a) 90
  (b) 95
  (c) 100
  (d) 110
  (e) 120

(7) Calculate the final quantity when a quantity decreases by 30% from 70.
  (a) 40
  (b) 45
  (c) 50
  (d) 49
  (e) 55

(8) Calculate the new value when an amount increases by 10% from $120.

(a) 110
(b) 115
(c) 120
(d) 126
(e) 132

(9) Calculate the original value when an amount decreases by 25% to reach $150.
(a) 200
(b) 180
(c) 175
(d) 160
(e) 150

BODMAS (Order of Operations)

(1) Simplify the expression: $3 + (4 \times 2) - \frac{5}{2}$.
(a) 7
(b) 8
(c) 8.5
(d) 9
(e) 10

(2) Evaluate the expression: $2 \times (6 + 3) - \frac{4}{2}$.
(a) 12
(b) 16
(c) 18
(d) 20
(e) 24

(3) Simplify the expression: $5 + 2 \times (8 - 3)$.
(a) 15
(b) 18
(c) 20
(d) 12
(e) 25

(4) Solve the expression: $4 \times \left(\frac{6}{2}\right) + 5 - 3$.
(a) 10
(b) 12
(c) 13
(d) 14
(e) 15

(5) Simplify the expression: $2 + (3 \times 4) - \frac{6}{2}$.
(a) 8
(b) 9
(c) 10
(d) 12
(e) 11

(6) Evaluate the expression: $(5 + 3) \times 2 - \frac{4}{2}$.
(a) 14
(b) 15
(c) 16
(d) 12
(e) 18

(7) Simplify the expression: $3 \times (4 + 2)/2$.
(a) 6
(b) 9
(c) 12
(d) 15
(e) 18

(8) Solve the expression: $2 + 3 \times \left(\frac{6}{3}\right)$.
(a) 5
(b) 6
(c) 8
(d) 9
(e) 10

(9) Simplify the expression: $4 + (5 \times 2) - \frac{3}{3}$.
(a) 10
(b) 11
(c) 12
(d) 14
(e) 13

(10) Simplify the expression: $(6 + 2) \times \frac{3}{2} - 5$.
(a) 7
(b) 8
(c) 9
(d) 10
(e) 6

Category	Likert Scale Questions
Performance	(PF1) Biblo does its task well (PF2) Biblo does not hinder me (PF3) I am capable of succeeding with Biblo
User's Engagement	(UE1) I was concentrated during the interaction with Biblo (UE2) The interaction captured my attention (UE3) I was alert during the interaction with Biblo
User's Trust	(UT1) Biblo always gives good advice (UT2) Biblo acts truthfully (UT3) I can rely on Biblo
User-Agent Alliance	(UAL1) Biblo and I have a strategic alliance (UAL2) Collaborating with Biblo is like a joint venture (UAL3) Biblo joins me for mutual benefit (UAL4) Biblo can collaborate in a productive way (UAL5) Biblo and I are in sync with each other (UAL6) Biblo understands me
Agent's Coherence	(AC1) [R] Biblo's behavior does not make sense (AC2) [R] Biblo's behavior is irrational (AC3) [R] Biblo is inconsistent (AC4) [R] Biblo appears confused
Agent's Intentionality	(AI1) Biblo acts intentionally (AI2) Biblo knows what it is doing (AI3) [R] Biblo has no clue of what it is doing (AI4) Biblo can make its own decision

Table 2: Likert Scale Questions by Category