Challenges in Exploiting Conversational Memory in Human-Agent Interaction

Socially Interactive Agents Track

Joana Campos
Disney Research
Carnegie Mellon University
Pittsburgh, PA, USA
joanacfcampos@cmu.edu

James Kennedy
Disney Research
Los Angeles, CA, USA
james.kennedy@disneyresearch.com

Jill F. Lehman
Disney Research
Carnegie Mellon University
Pittsburgh, PA, USA
jfl@andrew.cmu.edu

ABSTRACT

In human interactions, language is used to project and maintain a social identity over time. The way people speak with others and revisit language across repeated interactions helps to create rapport and develop a feeling of coordination between conversational partners. Memory of past conversations is the main mechanism that allows us to exploit and explore ways of speaking, given knowledge acquired in previous encounters. As such, we introduce an agent that uses its conversational memory to revisit shared history with users to maintain a coherent social relationship over time. In this paper, we describe the dialog management mechanisms to achieve these goals when applied to a robot that engages in social chit-chat. In a study lasting 14 days with 28 users, totaling 474 interactions, we find that it is difficult to leverage the shared history with individual users and to also accommodate to expected conversational coordination patterns. We discuss the implications of this finding for long-term human-agent interaction. In particular, we highlight the importance of topic modeling and signaling explicit recall of previous episodes. Moreover, the way that users contribute to interactions requires additional adaptation, indicating a significant challenge for language interaction designers.

KEYWORDS

Human-Robot Interaction; Conversational Agent; Memory

ACM Reference Format:

Joana Campos, James Kennedy, and Jill F. Lehman. 2018. Challenges in Exploiting Conversational Memory in Human-Agent Interaction. In *Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), Stockholm, Sweden, July 10–15, 2018,* IFAAMAS, 9 pages.

1 INTRODUCTION

Recent advances in the development of *Embodied Conversational Agents* (ECAs), whether as robots or virtual entities, has been propelled by technological progress in areas like natural language processing (NLP), graphics, and animation, to name a few. The rapidly growing field of NLP in particular has created opportunities for better spoken dialogue interactions between humans and ECAs, mostly in task-oriented domains [15]. Still, many challenges remain: not only does a dialog system need to find the appropriate thing to say in the interaction context, but it must also use relevant and meaningful conversational language to establish and maintain a coherent social relationship over time.

In human interactions, an important mechanism that helps us hold conversations, develop rapport, and maintain long-term relationships is our *memory* [23]. Without necessarily noticing, we constantly predicate our actions in the current situation to revisit and re-establish connections to our mutual past. Similarly, in ECAs, episodic memory architectures are believed to be essential as they aim to reflect the agent's identity and social awareness [17].

In this paper, we seek to extend prior incremental learning approaches [16]. We describe the dialog management mechanisms of a robot that engages in social chit-chat and uses the conversation history to revisit mutual past with the user to maintain a coherent social relationship over time. This approach goes beyond the use of event-related information and explores identity processes in a dyadic conversation, i.e., the agent uses a strategy to accommodate to the user's language and typical information-flow in previous encounters. This is a well known phenomenon in human-human communication that intends to create, maintain or decrease social distance in an interaction [6, 13].

The main focus of the research is to explore the challenges associated with persistence in a human-robot relationship over time. We are concerned in particular with how to capture what is revealed in a language-based interaction such that it can be used efficiently and efficaciously in future interactions with the same individual. To explore this topic, we use a purely statistical approach for dialog management, in a chat-oriented domain, in a study lasting 14 days with 28 users, totaling 474 interactions. We find that it is difficult to leverage the shared history with individual users, despite the reasonably high number of interactions. We draw from the data we have collected to highlight particular challenges for the language-based interaction community and suggest possible ways in which they may be addressed.

2 RELATED WORK

Conversational agents are becoming increasingly commonplace in everyday interactions. People can now use spoken language to execute commands with various appliances in their homes using goal-based dialog interfaces (e.g., Siri, Alexa or Cortana). Such interactions have become possible due to technological improvements in language-understanding systems and natural language tools in general [15]. These systems take advantage of large amounts of data

Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), M. Dastani, G. Sukthankar, E. André, S. Koenig (eds.), July 10–15, 2018, Stockholm, Sweden.

and are deployed to perform various tasks such as personal assistance [19], customer support [24, 34, 37], tutoring [36], training [33] and coaching [12]. A variety of machine learning approaches are a central element in such dialog systems, enabling the creation of agents that are more adaptive over time and less dependent on rules to generate responses (e.g., ELIZA [32] and ALICE [30]). Although current approaches offer an effective means of communication, they do not generalize well across domains as they are trained to be aligned to the task's demands. Therefore, other methods to generate responses for more generic systems have been pursued, such as machine translation [26], retrieval-based response selection [4], sequence-to-sequence models [10, 28, 29] and reinforcement learning for policy selection [20, 35]. To cite an example, Li et al. [20] modify Seq2seq to use RL, optimizing future rewards by capturing global conversation properties. They intend to learn an optimal policy for the conversation flow. Note that these algorithms require massive amounts of training data to perform well, which is not available for all domains (and is difficult to acquire).

Furthermore, designing Embodied Conversational Agents (ECAs) requires more than a conversational interface because such agents are intended to be lifelike and believable in their actions and reactions when interacting with humans over time [9]. Generating a natural conversation that is engaging over time puts many demands on system design, for example, scalability, adaptability, and a realistic management of dialog [27], even when domain knowledge is not available or known. Moreover, language use is highly correlated with large-scale social variables (e.g., age, gender or culture) [6] that should be addressed to achieve lifelike interaction.

One step toward more realistic dialog management could be achieved through effective communication and conversational coordination. In human-human interactions, the way people adjust their communicative styles can be explained by *Communication Accommodation Theory* (CAT) [13]. This accommodation can occur at many levels, from people adjusting the way they dress, to their accent, or the back-channeling behavior they use. The focus of our work is on language behavior and the discourse structure of the interaction between the user and the agent. CAT identifies two core concepts of accommodative behaviors: *convergence* and *divergence*. Converging to a common communicative style is associated with improvement in communication, lower uncertainty, higher predictability and mutual understanding. Divergence (non-accommodation) in communicative style conveys the opposite and is associated with expectation violation and less satisfaction in communication [13]. In this work we propose to take advantage of the conversational history between agent and human to improve communication and conversational coordination.

Current episodic memory architectures for ECAs are designed to store and manage relevant episodes collected in prior interactions for use in future dialogs, according to specific interaction goals. For instance, SARA [21] acts as a personal assistant at an event (e.g., conference) and uses task and social history with the users to make recommendations about conference sessions. Another example is MAY [7], which allows the user to talk about their experiences and use event-specific knowledge, linked to important moments of one's life, to match small talk templates in subsequent interactions. In both systems, event recollection is used to support the underlying task and build rapport over time. Elvir et al. [11] pursue an idea of *conversational memory* and advocate that conversational systems do not need to remember everything about a conversation, but instead only get the gist of what past conversations were about.

With the same intent, we take a different approach. We explore the development of a *conversational memory*, in an open chit-chat domain, that keeps track of the personal characteristics and language of a conversational partner. This relates to the work of Kennedy et al. [16] for instance, which describes an embodied agent that self-authors its own dialog for social chit-chat. It does so by incrementally building its knowledge base from its face-to-face interactions and a crowdworking pipeline. This is an extension of an earlier system [18] that used generation of narrative descriptions of future task situations to elicit dialog lines from crowdworkers. Building on [16] and [18], we apply novel techniques to incrementally build a language space of past interactions that enables the agent to revisit shared history and accommodate to the user's language, similarly to the phenomenon in human-human interactions.

3 RESEARCH QUESTIONS

Following from the related work, we are introducing a novel mechanism through which we aim to explore the use of a shared language history in interactions with people. It is desirable to understand both the **opportunity** for leveraging this history, and the **effect** of doing so. These two considerations drive our research questions for the study presented later in the paper.

RQ1 Is it possible to appropriately leverage conversational memory in continuing chat with users?

This question seeks to address whether or not it is possible to pursue the goals that an agent may have if it is focused on revisiting (or avoiding) a shared history with a user. Although the agent has a strategy for dialog selection, the user is an equal partner in any conversation, with simultaneous control over the interaction. It is not necessarily apparent that the agent will always be able to follow its strategy when seeking to leverage prior history.

RQ2 What effect does leveraging conversational memory have on an interaction?

Prior research suggests that language is predicated on prior history [23] and that this influences social distance between interactants [13]. Also, it has been shown in previous research that agents with memory are perceived as more competent [3]. We wish to explore whether these kinds of effects transfer to interactions between humans and robots, or hold true in our approach, motivating *b)* and *c)* below. As such, this research question can be broken down into more specific sub-questions:

- **a)** Do users perceive when a robot revisits language from their shared history?
- **b)** Do users experience increased satisfaction with the communication when a robot revisits shared language?
- **c)** Is the competency of the agent dependent upon use of the shared history?

4 ROBOT BEHAVIOR AND SYSTEM DESIGN

To explore the research questions, a robot was employed as a conversational agent in an office space. The sole task of the robot was to engage in social chit-chat, with the agent's internal goals for

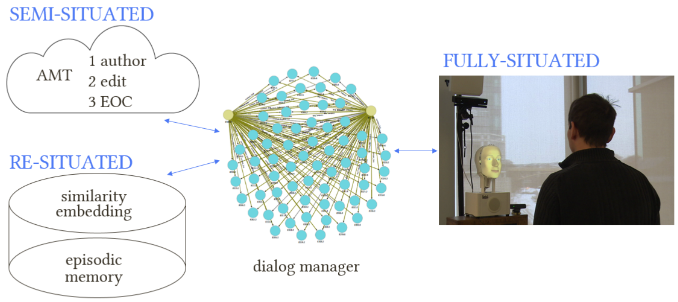

Figure 1: System diagram identifying the source of the three implemented learning mechanisms. A user interacting during the study with the robot is also depicted.

the conversation being driven by the research questions for the study. Specifically, this is a variation in whether the robot aims to revisit or avoid shared language history with users. Sec. 4.1 outlines the learning mechanisms derived from prior work [16], with the remainder of this section dedicated to the novel contribution here related to the agent's memory and decision strategy.

A Furhat retro-projected robot head [1] was used, with a male face. All users were provided with RFID tags for identification. The robot was supported with a Microsoft Kinect v2 for tracking head positions and a Logitech C920 camera for speech input. Upon approaching the robot, the tracked body would be matched with the RFID tag and the robot would have a 50/50 chance of initiating the conversation, or waiting for the user to initiate the conversation. If the user left the area around the robot and lost tracking then the interaction would end. The robot was provided with some baseline behavior to track users through moving its head for larger distances, or eyes for shorter distances. Lip sync is also automated on the platform.

4.1 Learning Mechanisms

The robot behavior and overall system framework was designed to encompass different mechanisms for language learning and use, incorporating various levels of situatedness. Different mechanisms are employed for learning depending on the existing knowledge in the system and the actions of both users and the robot during interactions; details can be seen in Sec. 4.2. The learning scenarios include semi-situated, re-situated and fully situated learning, outlined below:

Fully-situated learning. The robot can learn utterances from users during face-to-face interactions. Users are fully situated in the interaction: they have a complete representation of the world and conversational history. To eliminate speech recognition errors, these utterances are validated through the *edit* phase of the Amazon Mechanical Turk (AMT) pipeline, described next. If positively validated, the transcription of the user's utterance is added to the dialogue graph. This is done every time the user says something that is not similar to something already stored in the dialogue graph.

Semi-situated learning. Amazon Mechanical Turk is used as a crowd-based semi-situated learning environment and for speech recognition validation of the fully situated learning. Short narratives are provided to the crowd workers to expose some of the interaction state, along with relevant dialog history. Tasks will either request that workers contribute a continuation (*author* phase), judge a continuation (*edit* phase), or judge whether an utterance is an appropriate end of a conversation (*EOC* phase). All continuations (regardless of the learning origin) are validated using the *edit* and *EOC* phases. Learned utterances are removed from the system unless they meet a minimum criteria of being judged as a socially appropriate response that makes sense given the interaction state and history.

Re-situated learning. By using a word embedding space it is possible to examine semantic similarities between utterances. The word2vec embedding space used here utilizes soap opera scripts to reflect the social nature of utterances expected during interactions [22]. This provides a basis for using utterances that have not been validated previously in interaction states, or in response to things that do not currently have a continuation, but have a semantic closeness. Episodic memory of all interactions is used in combination with the semi-situated learning pipeline to manipulate how likely re-situated learning is to take place. This is a corrective function to prevent repeatedly making incorrect re-situations of language (further details follow in the next section).

In learning and execution of behavior over time, the combination of learning mechanisms is designed to take advantage of regularities in four different language distributions: those of the current user as revealed by episodic memory, of the community that interacts with the robot and each other, of a larger human community that experiences the situation narratively rather than experientially, and of an even larger community that has described similar kinds of events. A challenge is then presented in making decisions when faced with choosing and combining language from across these distributions.

4.2 Decision Processes

A dialog manager controls the learning mechanisms and stores all of the utterances in a graph database, along with a complete history of conversations with each user. This allows the robot to traverse the graph and its history when selecting an utterance to say. The dialog manager is also used when selecting utterances to test using the semi-situated pipeline.

The traversal process is taken each turn based on what the user has said (or context of the conversation when the robot takes the

Exploration		Exploitation	
U	Tell me what kind of movie do you like.	R	You ready for today?
R	I like crime thrillers	U	Yeah I am, are you ready for today?
U	What's your favorite movie?	R	I've never been more prepared for a presentation
R	Batman	U	What's your presentation on?
U	What movie are you going to see?	R	You ready for today?
R	I'm not sure yet	U	I'm ready for today.
U	What's your favorite movie?	R	I've never been more prepared for a presentation
R	I don't like movies	U	What's your presentation?

Figure 2: Examples of exploiting and exploring language history (dialog excerpts taken from study presented later in this paper). 'U' represents user utterances, 'R' represents robot/agent utterances. The bottom excerpts represent later timesteps, where the agent response is contingent on the action taken in the earlier timestep (top excerpts).

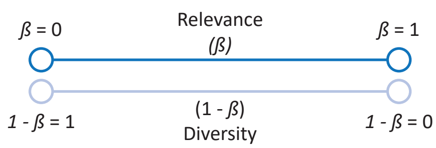

Figure 3: The *Diversity-Relevance* continuum. For high values of β there is less *diversity*.

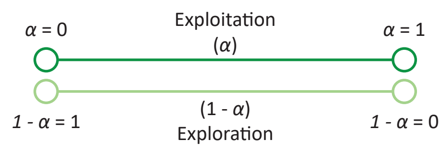

Figure 4: The *Exploration-Exploitation* continuum. Higher values of α prioritize exploitation over exploration.

first turn) and returns a set of *candidate answers*. If this set is not empty, an utterance is selected according to the agent goals (described in Sec. 4.3) and the algorithm in Sec. 4.4. The set of *candidate answers* includes any possible continuation to what the user previously said across all utterance-space (U). This implies that dialog utterances can be 'borrowed' across contexts within the graph as a form of generalization. This is done by creating a temporary link between two vertices in the dialog graph, connecting what the user said with a selected dialog utterance. This connection is only made permanent if three AMT workers rate it as plausible. This process does not increase the agent's repertoire of utterances but allows the agent to expand its ability to continue a conversation.

Finally, if the agent didn't find anything similar to what the user said or if it has but did not have an answer for it, the agent ends the conversation with "Oops, gotta go!" and generates an AMT task to avoid this failure point in the future.

4.3 Agent's Internal Goals

The goals of an agent are a principled way of defining preferences. We use the same formalism for defining goals as in [2], which characterizes a goal $g_i \in G$ in terms of three parameters: *value* ($v_i \in \mathbb{R}$), which denotes the importance of a goal; *priority* ($r_i^t \in \mathbb{N}$) and *degree of achievement* ($d_i^t \in [0,1]$). Both priority and degree of achievement are time dependent, as they may differ between time-steps. The value of a goal is fixed and does not change over time. These three concepts contribute to an evaluation of how an utterance choice fulfills the agent's goals (more details in Sec. 4.4).

Let U = $\{u_1, ..., u_n\}$ represent a finite set of possible utterances (where n is the total number of nodes in the knowledge base) and $a \in \{K, H\}$, represent the Robot and the Human, respectively. Let $X^{T_n}_{R\leftrightarrow H}$ be a set of dialogue turns ($x^{t_1}_{R\rightarrow H}, x^{t_2}_{H\rightarrow R}, ..., x^{t_i}_{a\rightarrow a}$) between R and H, where i is the total number of turns. Furthermore, each $x^{t_i}_{a\leftrightarrow a}$ is an utterance from the set U. At each time step t the agent needs to select an action $x^t_{R\rightarrow H}$ given the most recent recent pair of turns ($x^{t_i-2}_{R\rightarrow H}$, $x^{t_i-1}_{H\rightarrow R}$). The implemented selection mechanism, rooted in the agent's goals, considers each node in the graph as a possible action to take. The goals are represented according to four concepts: *diversity*, *relevance*, *novelty*, *past experience* [5].

*g_1**: Relevance**.* A node is relevant if it is appropriate to use given the pair ($x^{t_i-2}_{k\rightarrow h}, x^{t_i-1}_{h\rightarrow k}$) and if it belongs/was used in similar context.

*g_2**: Diversity**.* Diversity can be understood as the opponent force to relevance. Consider β, a control parameter that tunes the trade-off between *diversity* and *relevance* (see Fig. 3). This means that the higher the diversity, the lower the relevance.

*g_3**: Past Experience (or Exploitation)**.* This refers to the motivation to re-visit the users' past history, including language use and information-flow.

*g_4**: Novelty (or Exploration)**.* The novelty of a node or a set of nodes is defined by how far, or diverse, an item is from the user's past experiences. If we consider α as a control parameter (see Fig. 4), then in a pure exploration strategy ($\alpha = 0$) the agent will search everywhere, except in the user history.

4.4 Utterance Selection

This agent-based solution can be approached as a filtering system that seeks to find the most suitable item from a large set of nodes, taking into consideration relevance and the goals of the system (which may include explore different topics or exploit a past interaction). The utterance selection mechanism (illustrated in Fig. 5) chooses the utterance u_i from the set of available utterances U, such that it maximizes the degree of achievement of the agent's goals. We use the same rationale as in [2]. The evaluation of how good an utterance depends on the three defining elements of a goal (*value* - v_i, *priority* - r_i^t and *degree of achievement* - d_i^t). Although the mechanism is generic, it requires the definition of a function f to compute the impact of each action given the current state of

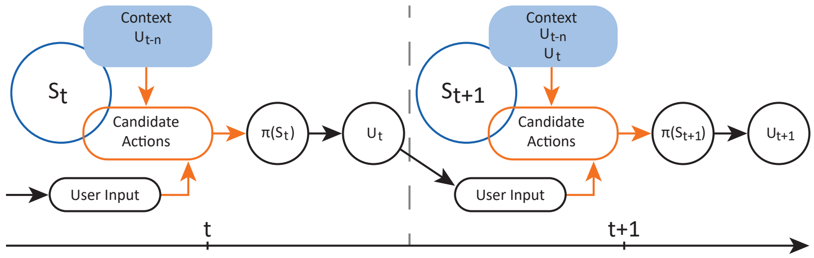

Figure 5: Utterance selection procedure. At each time step t, the agent evaluates the user input and current state, S_t, against its policy (π) for utterance selection, defined by its goals. The agent selects the utterance (u_t) that maximizes its goals at time t.

the world, that is, how an utterance changes d_i^t for each goal. This function is domain dependent and should consider the current state of interaction.

At each time step, the agent acts to increase the degree of achievement of its goals. Therefore the aforementioned elements are combined to create a valuation function that expresses the agent's expected utility at time t ($\hat{u}^t$):

$$\hat{u}^t = \sum_{g_i \in G} d_i^t \cdot q(v_i \cdot r_i^t)$$

To maximize $\hat{u}^t$ the agent has to possess some domain knowledge about the available set of actions in the world and how those will affect the degree of achievement of the agent's goals. Therefore we define a function $f(d_i^t, a|\sigma^t)$ that calculates the impact of a on d_i^t given the current state of the world, σ^t. The action that maximizes the expected increase in degree of achievement of the agent's goals weighted by their relative importance and current priority is selected:

$$a_t \in argmax_{a \in A} \hat{u}(r^t, f(d_i^t, a|\sigma^t))$$

The *q-function*, $q(v_i \cdot r_i^t)$, defines how the value and priority of a goal interact. We use the same function as described in [2]: $q(v_i \cdot r_i^t) = v_i \cdot (r_i^t / r_{max})$, where $1 \leq r_i^t \leq r_{max}$. Other functions could be used, but this quantification is appropriate given the research questions here: the lower the priority, the lower the relative importance of a goal. This allow us to create a hierarchy of goals that establish that all active goals are considered for utterance selection, but some goals are more important than others. In the formalization of our system, *g1:Relevance* is the most important goal in this selection mechanism (goal with higher priority), meaning that the agent will prioritize language that is contextually more relevant according to the current state of the system.

For the experimental evaluation described in Sec. 5 we created two agents: one that values more *g3:Exploitation* and another that values *g4:Exploration*. In one extreme, we have an agent that prioritizes selecting language that has been used before in the interaction (i.e., from a shared history), whereas the other agent will always try to explore new language. This output is generated by the authored priority of the goals (r_i^t) and the valuation function applied to the set of candidate answers. Although this value could be changed over time depending on the evolution of the system where the agent was deployed, in this work the values were fixed.

5 EVALUATION

5.1 Participants

A total of 28 users (age M=29.32, SD=8.9), 13 female and 15 male, participated in a study in an office space over a period of 3 weeks (14 days total), totaling 474 interactions with the robot. The study was conducted with IRB approval. Participants provided informed consent and were paid for their time. Video and audio were recorded for post-hoc analysis. Users wore RFID tags to enable recognition by the agent.

The number of participants was not the same each week: 26 users participated in Week 1 (interactions M=5.12, SD=3.33; most turns for each user M=10.34, SD=6.50), 16 users in Week 2 (interactions M=10.37, SD=6.00; most turns for each user M=11.31, SD=3.82) and 20 in Week 3 (interactions M=8.75, SD=4.70; most turns for each user M=13.65, SD=5.34).

5.2 Procedure

We conducted a between-subjects experiment to explore the research questions. Users were split into two groups: the *exploitation* condition and the *exploration* condition. In both conditions, the agent's behavior is driven by its goals. In the former, the agent was configured to select dialogue utterances that maximize *the exploitation of their past history*. In the latter, the agent's maximization function selects utterances that intend *to explore anything but the past history* between the agent and the user. Note that in both conditions the agent prioritizes relevant utterances, i.e., utterances that the agent is more certain to be appropriate for that interaction context. At the end of each week, participants were requested to answer a brief on-line questionnaire about their interactions that week and the agent's behavior in general.

Users were not aware of the experimental manipulation and were instructed to approach the agent as if it were simply a person in their workplace that they knew only casually. We asked them to adopt a benevolent attitude, i.e., to not try to make it fail. Additionally, we instructed users to interact with the robot at least 3 times a day.

5.3 Measures

In this experimental setting both the robot and the user have control over the interaction, which may hamper or facilitate the robot's **opportunity** for leveraging prior history (**RQ1**). To analyze the resulting interaction, we define two characteristics to capture whether the robot employed its intended strategy (of *explore* or *exploit*):

Match. The robot's utterance is classified as a *match* when the robot followed its intended strategy. For each week, for each user, the overall percentage of *match* utterances was calculated. If *match* percentage is equal to or higher than the average, the robot behaved according to its design.

Mismatch. When the *match* percentage is lower than average we consider that the robot employed a strategy different from its initial design. That is, if the robot was designed to exploit, but the use of that strategy was below average we consider that overall, the robot actually explored.

Using the same rationale, it is possible to categorize users by their alignment to the robot strategy. This allow us to characterize *(robot, user)* pairs as *convergent* or *divergent*, i.e., interacting parties that used or did not use the same strategy, respectively (this is motivated by CAT in Sec. 2). When the *(robot, user)* pair diverges in their communication style, **tension** exists in the interaction. This corresponds to situations where the robot is trying to revisit a past conversation path but the user is not willing to[1].

To test the **effect** of the manipulation (**RQ2**), at the end of each week, participants responded to an adaptation of the *Interpersonal Communication Satisfaction Inventory* (ICSI) [14] that was designed to assess interpersonal communication satisfaction (**RQ2b**).[2] Additionally, one dimension – *competence* – of the validated RoSAS questionnaire [8] was included, which was designed to assess user's social perception of robots (**RQ2c**). Users were asked to classify how closely the words *capable*, *responsive*, *interactive*, *reliable*, *competent*, and *knowledgeable* were associated with the robot on a scale from 1 (*not all associated*) to 7 (*definitely associated*). The results from the questionnaires, completed at the end of each week, represent an assessment of the communicative experience over multiple interactions with the robot to address.

Additionally, at the end of the third week, we asked users to classify the sentences: *"I felt I had repetitive conversations with the robot"*; and *"I felt I talked about the same things with the robot"*, using a scale from 1 (never) to 5 (always) to classify whether they perceived that the robot revisited language from their shared history (**RQ2a**). This was only done at the end so that subjects did not focus on this aspect whilst the experiment was running (to prevent a possible confound).

All interactions were transcribed in order to assess the noise created by the Automatic Speech Recognition system (cloud-based Microsoft Bing). There is a 15% word error rate (WER) considering all dialog turns, indicating a favorable performance.

6 RESULTS AND ANALYSIS

6.1 Design conditions and questionnaire data

The two experimental conditions, ***exploitation*** and ***exploration***, correspond to an agent that is only concerned in exploiting its past history and an agent that wants to avoid the past history, respectively. Although in both conditions each decision of the agent is taken to maximize the achievement of its goals, the agent is not able to apply a pure strategy over time (Fig. 6; **RQ1**). There are several factors affecting the robot's behavior. First, the priority of the agent is to select context-relevant utterances, meaning that the agent may select something it knows will work in the current context instead of selecting something purely because it is in the user's past history, i.e., making sense is prioritized over saying something from the history. Second, the exploitative agent can only start behaving in the way it was designed after the first interaction, incrementally. Finally, the agent is not the only force in the interactions and the user can continuously push the agent way from (or toward) its path. This may create a situation where the agent has no other option but to do something contrary to its goals.

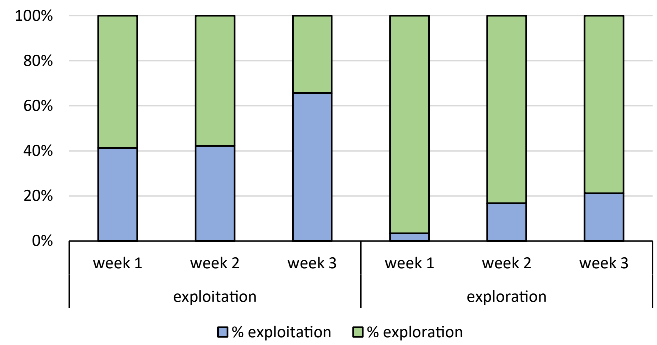

Figure 6: Percentage of use of both strategies by condition. The percentage of strategy use is calculated on a per turn basis across all users. During the first two weeks, the exploitation condition agent did not have the opportunity to follow the intended strategy the majority of the time.

To test the differences between the two conditions with regard to users' responses in the questionnaires, we conducted Mann-Whitney tests across the data. The results indicate that there are no significant differences in the users' satisfaction (**RQ2b**) or how users perceive the agent's competence, given the way the robot was designed to behave (**RQ2c**). Additionally, there was no difference between the conditions regarding user perception of the use of history (**RQ2a**). It should be noted that all of these questions are entangled with the opportunity for the agent to meet its goals. These issues will be returned to in the discussion (Sec. 7).

6.2 Dynamic adaptation

Our goal was to compare the users' perceptions of the interactions they had, given the condition to which they were assigned. However, users applied their own strategies as well. As a result, it is possible to view the data in terms of interactions where the parties mostly ***converge*** or ***diverge*** in their communication style, as outlined in Sec. 5.3. This allows separation of the users into two groups: users that *match* the robot's behavior and users that did not (*mismatch*). On this basis, it is possible to categorize users by their alignment to the robot strategy (*Robot strategy, User strategy*) summarized in Table 1. As a result of the behavior of the agent, participants adapted their behavior as well, generating different types of experiences for themselves - more convergent, when the same strategy is applied, or more divergent, when they apply different

[1]The opposite is also true.

[2]Participants answered 11 questions (out of 16) of the original questionnaire. Questions 1, 2, 7, 12 and 15 were excluded since they did not make sense in the created scenario. As a result, we were not able to compute the communication score that is the usual output of this validated scale.

		User Strategy					
		Week 1		Week 2		Week 3	
		A	*B*	*A*	*B*	*A*	*B*
Robot	*Exploitation (A)*	0	14	2	9	5	7
	Exploration (B)	0	12	0	5	1	7

Table 1: Summary of the behavior of the dyad each week. The value of each cell is the number of users in that situation. A = Exploitation; B = Exploration.

strategies. Despite this alternate way of framing the interactions, there are no significant differences in the questionnaire data when comparing the *convergent* versus the *divergent pairs*. Based on prior literature [13], we anticipated that people that converge may classify their interaction with the robot more positively than those that diverge, but this does not seem to be the case.

Focusing on *mismatch*, which is a force that moves the robot way from its dialogue trajectory and creates a *tension* in the conversation, we looked at the pairs (Robot, User) in each interaction. In particular, we inspected pairs of utterances with the pattern (*match*, *mismatch*), i.e., situations where the robot is trying to revisit a past conversation path but the user is not willing to. The number of these occurrences increases from week to week, as may be expected. The fact that the agent exploits more appears to motivate the user to try to change the conversation path. The average percentage of *tension* in dialogue grows from 25% in week 1 to 36% in week 3. Interestingly, this trend is not observed in the exploration condition (average occurrence of 10% in week 3), with the exception of one user who tries to have the same conversation but the robot is interested in talking about something else.

From the questionnaire data, participants that experienced more *tension* in the interaction agreed more with questions Q7[3] ($U = 6.5, z - 2.189, p = .029$) and Q9[4] ($U = 7.5, z - 2.079, p = .042$) of the ICSI scale, than those that did not experience as much tension (in the exploitation condition, in Week 1). In Weeks 2 and 3 there are not enough answers in the same condition for a fair comparison. Unsurprisingly, this raises attention to the fact that a purely exploitative system is not ideal. When the user and robot mostly exploited previous history, this increases repetition.

6.3 Length of conversation

The initial design of the experiment did not produce significant differences in the experience of the user, nor did the resulting adaptation process. It seems that more things happen in the interaction that affect the way users perceive the agent and the quality of the communication. In fact, there are strong correlations between the greatest number of turns a user achieves (each week) and their perception of the agent as knowledgeable, capable, competent and interactive. Additionally, the maximum number of turns also correlates with the users satisfaction and ability to say what one wants. Table 2 details these findings.

[3]"I felt I could talk about anything with the other person."
[4]"The conversation flowed smoothly."

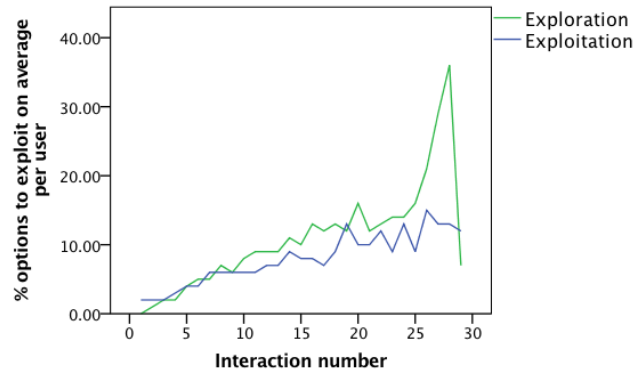

Figure 7: Percentage of options to exploit over time, calculated on a per turn basis across all users.

	Max number of turns		
	Week 1	Week 2	Week 3
Q4[1]	$r_s = -0.473$ ($p = .020$)	ns.	$r_s = -0.532$ ($p = .041$)
Q5[2]	ns.	ns.	$r_s = -0.619$ ($p = .014$)
Q8[3]	$r_s = -0.453$ ($p = .026$)	ns.	ns.
Q10[4]	$r_s = -0.412$ ($p = .045$)	ns.	ns.
Knowledgeable	$r_s = 0.440$ ($p = .032$)	ns.	ns.
Competent	ns.	$r_s = 0.606$ ($p = .037$)	ns.
Interactive	ns.	ns.	$r_s = 0.555$ ($p = .032$)
Capable	ns.	ns.	ns.

[1] Q4 : "I was very satisfied with the conversation"
[2] Q5 : "The other person expressed a lot of interest in what I had to say."
[3] Q8 : "We each got to say what we wanted"
[4] Q10 : "The other person often said things that added little to the conversation"

Table 2: Spearman Correlations between questionnaire data and the max. number of turns a user achieved

6.4 Availability of language to exploit

As discussed above, the agent's ability to exploit previous paths of conversation will depend on the agent's behavior combined with the user's behavior. Therefore, if the user always says the same thing in the exploitative condition that will make the agent repeat itself several times. While this may be relevant in situations where the agent must be coherent, it does not increase the range of topics a (purely) exploitative agent is able to talk about. One of the reasons is that in this experimental setting, the agent is not intentionally pursuing topics of conversation, in contrast to systems like [21] or [7]. On the other extreme, the exploration agent creates more opportunities to exploit despite never taking advantage of that language space. Fig. 7 illustrates how the language space grows as a function of the strategy used. This is calculated by considering the number of options to exploit at each decision point, balanced by the number of options available. This suggests that the amount of growth that occurs in three weeks is not enough to explore the dynamics that are present in the system, given the learning rate.

7 DISCUSSION

It is a main premise in this work that having mechanisms to build a *conversational memory* allows the agent to revisit a shared past with the user. This enables adoption of the user's language and accommodation to the usual information-flow in previous encounters. The architectural backbone of our system leverage the learning mechanisms to reuse language according to social and contextual variables. We intended to understand both the **opportunity** (**RQ1**) for leveraging shared history and the **effect** (**RQ2**) of doing so. We found no differences between the designed conditions and that was due to several factors, which reflect significant interaction challenges (IC) that language interaction designers should address.

IC1 Users dynamically adapt to the interaction and have their own strategies.

We find that the scenario created, where we expose users to two conditions, ***exploitation*** of past history vs. ***exploration*** of new language, generated a dynamic adaptation process in the interaction. Users also employed one of the two strategies, generating different kinds of experiences than those initially designed, as summarized in Table 1. In particular, when the user constantly applied a strategy contrary to the robot's design strategy, that created *tension* in dialog movement. Users in the exploitation condition that experienced more tension reported that the conversation flowed smoothly and they felt they could talk about anything with the other person. A possible reason is that the user was able to redirect the conversation to a topic that pleased them more compared to those in the exploitation condition that revisited the same path several times.

IC2 Model conversation topic explicitly.

It is important to note that the notion of reusing history as presented in this paper is based on reasoning about clusters of semantically similar utterances. These clusters do not necessarily reflect topics of conversation, so there is a potential for the agent to repeat a topic without having the intent to do so. Typically, revisiting a topic of conversation is done to build rapport, or to add something new to a topic that was previously discussed [13]. Without a model of topic to reason about, it may be that even with more interactions and history exploitation opportunities, the experience is not positive for the user due to this repetition without novel contribution. If the system were able to reflect on its own history and knowledge structures, with a notion of topic that could be inspected, then it might be possible to better identify user interests. Such an approach would not necessarily require a deep understanding of semantics (beyond the topic level), but would allow an expanded concept of conversational goals. In turn, these conversational goals could be used to motivate learning – that is, learning can be focused on topics that the agent knows that users are interested in. This, in turn, would allow the agent to build on the shared content while avoiding repetition, and further engage the user in 'meaningful' social chat. By focusing the learning in this manner, the agent would be simultaneously optimizing resources (the time and cost of content generation). Although substantial progress has been made in topic modeling and segmentation, it is still an open research problem that is harder in face-to-face interactions due to characteristics of the language and the lack of training data [31].

IC3 Explicitly signal recall in a meaningful way.

Sec. 6.1 highlights a potential issue in whether people perceive the exploitation of language that the robot attempts. This is complicated by the robot having limited chances to follow its goals of re-using prior history (Sec. 6.4). If the robot always pursues the same topics or language when interacting with a user, it may not be clear that the robot is intentionally recalling a shared past. The user may simply believe that this is the only thing the robot knows about, or the only way it has to express something. It seems plausible that to fully realize the intent of the robot, people must be exposed to a variety of behavior so that they are aware it *can* do more, but is intentionally recalling shared history. Of course, by expressing a variety of behavior, this behavior then also becomes part of that shared history. Alternatively, the robot must learn some means of signaling to the user that the decision to revisit something from memory is an intentional decision. This could be through language that exposes the robot's internal goals, e.g., 'I know we talked about this the other day, but I want to talk about it again with you'. In a recent study, Richards and Bransky [25] found that recall of information about the user increases the user's enjoyment of interaction and agent believability. However, the critical aspect is to infer what is important to recall in future interactions and to signal that recall in a way that users perceive as meaningful.

8 CONCLUSION

Embodied Conversational Agents (ECAs) that intend to engage in long-term relationships with humans face several challenges when the novelty effect wears off. The main challenge is to find new ways to generate content and new behaviors for an agent that must interact in socially-appropriate ways repeatedly over an extended period of time. Within the possible solutions for maintaining user engagement is the integration of a memory mechanism. In this paper, we explored the creation and exploitation of a *conversational memory* over multiple interactions with several users. This was done by determining the next utterance to say based on the agents goals, which might include exploiting past history or exploring new areas. We find that revisiting conversational memory requires more than implicitly following previous conversation paths (**IC3**). Instead, mechanisms should be created to make the user aware of the agent's intention to revisit shared experience. Furthermore, taking advantage of a *conversational memory* is about having the ability to modify the agent's communication style based on idiosyncratic characteristics of the conversational partner [13]. For instance, being able to select to talk about cinema versus initiating a conversation about novels, because the agent knows the user's interests. This type of reasoning relies on the notion of topic which should be the focus of future work (**IC2**). Even the simple grouping of users' interests into explicit topics might allow the agent to optimize resources for learning. By re-framing its conversational goals the agent may decide when and how to expand specific topics based on the interaction needs (**IC1**).

REFERENCES

[1] Samer Al Moubayed, Jonas Beskow, Gabriel Skantze, and Björn Granström. 2012. Furhat: a back-projected human-like robot head for multiparty human-machine interaction. *Cognitive Behavioural Systems* (2012), 114–130.

[2] Dimitrios Antos and Avi Pfeffer. 2011. Using emotions to enhance decision-making. In *IJCAI Proceedings-International Joint Conference on Artificial Intelligence*, Vol. 22. 24.

[3] Ruth Aylett, Michael Kriegel, Iain Wallace, Elena Segura, Johanna Mercurio, and Stina Nylander. 2013. Memory and the design of migrating virtual agents. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*. International Foundation for Autonomous Agents and Multiagent Systems, 1311–1312.

[4] Rafael E Banchs and Haizhou Li. 2012. IRIS: a chat-oriented dialogue system based on the vector space model. In *Proceedings of the ACL 2012 System Demonstrations*. Association for Computational Linguistics, 37–42.

[5] Andrea Barraza-Urbina, Benjamin Heitmann, Conor Hayes, and Angela Carrillo Ramos. 2015. XPLODIV: An Exploitation-Exploration Aware Diversification Approach for Recommender Systems.. In *FLAIRS Conference*. 483–488.

[6] Leslie M Beebe and Howard Giles. 1984. Speech-accommodation theories: A discussion in terms of second-language acquisition. *International journal of the sociology of language* 1984, 46 (1984), 5–32.

[7] Joana Campos and Ana Paiva. 2010. MAY: My Memories Are Yours.. In *IVA*. Springer, 406–412.

[8] Colleen M Carpinella, Alisa B Wyman, Michael A Perez, and Steven J Stroessner. 2017. The Robotic Social Attributes Scale (RoSAS): Development and Validation. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 254–262.

[9] Justine Cassell. 2000. *Embodied conversational agents*. MIT press.

[10] Jesse Dodge, Andreea Gane, Xiang Zhang, Antoine Bordes, Sumit Chopra, Alexander Miller, Arthur Szlam, and Jason Weston. 2015. Evaluating prerequisite qualities for learning end-to-end dialog systems. *arXiv preprint arXiv:1511.06931* (2015).

[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.

[12] Juan Fasola and Maja Mataric. 2013. A socially assistive robot exercise coach for the elderly. *Journal of Human-Robot Interaction* 2, 2 (2013), 3–32.

[13] Howard Giles, Anthony Mulac, James J Bradac, and Patricia Johnson. 1987. Speech accommodation theory: The first decade and beyond. *Annals of the International Communication Association* 10, 1 (1987), 13–48.

[14] Michael L Hecht. 1978. The conceptualization and measurement of interpersonal communication satisfaction. *Human Communication Research* 4, 3 (1978), 253–264.

[15] Julia Hirschberg and Christopher D Manning. 2015. Advances in natural language processing. *Science* 349, 6245 (2015), 261–266.

[16] J. Kennedy, I. Leite, A. Pereira, M. Sun, B. Li, R. Jain, R. Cheng, E. Pincus, E. Carter, and J.F. Lehman. 2017. Learning and Reusing Dialog for Repeated Interactions with a Situated Social Agent. In *Proceedings of the International Conference on Intelligent Virtual Agents*.

[17] John E Laird and Nate Derbinsky. 2009. A year of episodic memory. *Ann Arbor* 1001 (2009), 48109–2121.

[18] I. Leite, A. Pereira, A. Funkhouser, B. Li, and J. F. Lehman. 2016. Semi-situated Learning of Verbal and Nonverbal Content for Repeated Human-robot Interaction. In *ICMI 2016*. ACM, New York, USA, 13–20.

[19] Oliver Lemon. 2012. Conversational interfaces. In *Data-Driven Methods for Adaptive Spoken Dialogue Systems*. Springer, 1–4.

[20] Jiwei Li, Will Monroe, Alan Ritter, Michel Galley, Jianfeng Gao, and Dan Jurafsky. 2016. Deep reinforcement learning for dialogue generation. *arXiv preprint arXiv:1606.01541* (2016).

[21] Yoichi Matsuyama, Arjun Bhardwaj, Ran Zhao, Oscar Romeo, Sushma Akoju, and Justine Cassell. 2016. Socially-Aware Animated Intelligent Personal Assistant Agent.. In *SIGDIAL Conference*. 224–227.

[22] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781* (2013).

[23] Katherine Nelson. 2003. Self and social functions: Individual autobiographical memory and collective narrative. *Memory* 11, 2 (2003), 125–136.

[24] Antoine Raux, Brian Langner, Dan Bohus, Alan W Black, and Maxine Eskenazi. 2005. Let's Go Public! Taking a spoken dialog system to the real world. In *Ninth European Conference on Speech Communication and Technology*.

[25] 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.

[26] Alan Ritter, Colin Cherry, and William B Dolan. 2011. Data-driven response generation in social media. In *Proceedings of the conference on empirical methods in natural language processing*. Association for Computational Linguistics, 583–593.

[27] Matthias Scheutz, Rehj Cantrell, and Paul Schermerhorn. 2011. Toward humanlike task-based dialogue processing for human robot interaction. *Ai Magazine* 32, 4 (2011), 77–84.

[28] Iulian Vlad Serban, Alessandro Sordoni, Yoshua Bengio, Aaron C Courville, and Joelle Pineau. 2016. Building End-To-End Dialogue Systems Using Generative Hierarchical Neural Network Models.. In *AAAI*. 3776–3784.

[29] Oriol Vinyals and Quoc Le. 2015. A neural conversational model. *arXiv preprint arXiv:1506.05869* (2015).

[30] Richard S Wallace. 2009. The anatomy of ALICE. *Parsing the Turing Test* (2009), 181–210.

[31] Liang Wang, Sujian Li, Yajuan Lv, and WANG Houfeng. 2017. Learning to Rank Semantic Coherence for Topic Segmentation. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. 1351–1355.

[32] Joseph Weizenbaum. 1966. ELIZAâĂŤa computer program for the study of natural language communication between man and machine. *Commun. ACM* 9, 1 (1966), 36–45.

[33] Atef Ben Youssef, Mathieu Chollet, Hazaël Jones, Nicolas Sabouret, Catherine Pelachaud, and Magalie Ochs. 2015. Towards a socially adaptive virtual agent. In *International Conference on Intelligent Virtual Agents*. Springer, 3–16.

[34] Zhou Yu, Dan Bohus, and Eric Horvitz. 2015. Incremental coordination: Attention-centric speech production in a physically situated conversational agent. In *Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*. 402–406.

[35] Zhou Yu, Ziyu Xu, Alan W Black, and Alexander I Rudnicky. 2016. Strategy and Policy Learning for Non-Task-Oriented Conversational Systems.. In *SIGDIAL Conference*. 404–412.

[36] John Zakos and Liesl Capper. 2008. CLIVE–an artificially intelligent chat robot for conversational language practice. In *Hellenic Conference on Artificial Intelligence*. Springer, 437–442.

[37] Victor Zue, Stephanie Seneff, Joseph Polifroni, Michael Phillips, Christine Pao, David Goodine, David Goddeau, and James Glass. 1994. PEGASUS: A spoken dialogue interface for on-line air travel planning. *Speech Communication* 15, 3-4 (1994), 331–340.