Proceedings of the 29th IEEE International Symposium
on Robot and Human Interactive Communication,
August 31 - September 4, 2020 - Virtual Conference

ThPIP.13

Improving Quality of Life with a Narrative Robot Companion: II – Creating Group Cohesion via Shared Narrative Experience*

Takahisa Uchida, Hiroshi Ishiguro, Peter Ford Dominey

Abstract— One of the most difficult things for social robots is to enter the realm of human social relations. Here, we exploit recent advances in natural language processing (NLP) that provide robots access to human experience, which can allow them to enter into human social relations. When people are put together in arbitrary conditions, as in a home for the elderly, it can be difficult for them to share experience. To address this, we use recent advances in NLP to allow the robot to discover shared narratives between such group members. Our narrative companion extends the human capability to make social narrative links for building group coherence through sharing experience.

Requirements are identified for a narrative companion to allow individuals within a group to focus their interactions on shared experiences and interests, to improve group coherence. The system should collect and organize members' experiences, and should discover semantic similarity between different members' experiences in order to create a group narrative. It should then accompany the group in their cohesion-enhancing experience of this narrative. Based on these requirements, and extending our previous work, we implement the V2.0 narrative companion on the Pepper robot.

The system is validated in a case study where participants provide 5 favorite photographs and short answers to questions. The Narrative Semantic Similarity Analysis System (NarSim) generates a meaningful trajectory of narrative linking people and events depicted in the photographs. With this, Pepper then accompanies the group and prompts group members to enrich the shared narrative, to further enhance the group pleasure and cohesion. Results are presented, and future applications for improved quality of life are discussed.

I. INTRODUCTION

Establishing and maintaining self-identity is central to human well-being [1, 2]. The self cannot exist without its relation to others [3], and thus the establishment of relations between individuals is crucial for well-being of members of the group [4]. We previously developed a narrative assistant that operates at the level of the individual, allowing one to store, retrieve and share autobiographical memories through narrative [5]. In the current research we extend our approach to a narrative assistant that creates cohesion between multiple individuals that form a group. Like individual selves, the group self is potentially fragile, and requires cohesion [4].

Group cohesion can result from sharing fun experiences, which can take a variety of forms including sharing new experiences, and shared recalling of past fun experiences [4].

When a group is constituted for external reasons, e.g. creating a scientific research team, an engineering design team; or placement of residents in a home for the elderly, it may be the case that the members have no knowledge of each other, and thus while they might have significant shared backgrounds and histories, these are unknown to them, and are potentially difficult to discover. To unblock such a situation, one could ask members of the group to provide a small set of favorite photographs from their past with short answers to questions about what, where, when, with who, and why they like the photo. From this collection of experience one could discover common shared interests, but how?
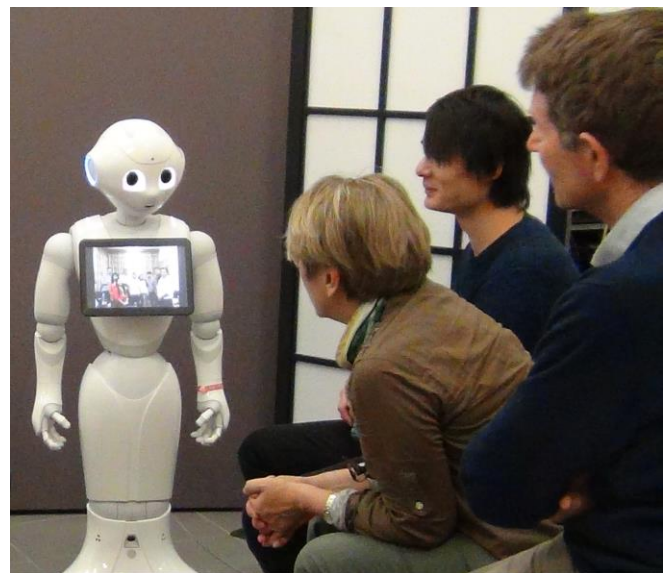

Figure 1. Pepper robot animating a group activity

In the current research we address this challenge, extending our previous research [5-8] to develop a robot narrative companion in a group context. We use distributed representation techniques developed in machine learning and natural langue understanding [9] in order to identify narrative relations between descriptions that different participants provide for their photographs. Distributed word representations are n-dimensional vector representations of words that are generated by extracting statistics on neighbor relations between words in large text training corpora [9]. The resulting vector representations have interesting semantic properties, such that vectors for words with similar meanings

T. Uchida and H. Ishiguro are with Ishiguro Lab, Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama Toyonaka Osaka 560-8531 Japan (email: uchida.takahisa , ishiguro @irl.sys.es.osaka-u.ac.jp)

P.F. Dominey is with INSERM UMR1093-CAPS, Université Bourgogne Franche-Comté, UFR des Sciences du Sport, F-21000 Dijon, France (email: peter.domioney@inserm.fr).

will be similar. Recent research has successfully extended these representation properties from word to text levels [10], and further research indicates that these distributed vector representations of stories can be related to distributed neural activity in the human brain while people read these stories [11]. We believe that these semantic processing capabilities can be integrated into narrative assistant robots.

The value and feasibility of robots as partners in narrative interaction has been demonstrated in child development of narrative skills [12]. Robots are now being introduced in the care of elderly with dementia to increase quality of life and reduce depression [13]. The potential for these robots is promising, and one need that has been identified is to personalize the robots to individual participants [14].

The main contributions of this work are: 1) We use word embeddings to identify semantic similarity relations between photo descriptions from different participants. 2) We use these automatically identified relations to create a multimodal dynamic group narrative that allows the members to discover, share, and exchange about these relations between them which were previously unknown. 3) This novel mechanism can help to establish meaningful personal relationships that can contribute to group coherence.

II. METHODS A: SYSTEM REQUIREMENTS

To build the companion, we take a cognitive systems engineering approach [15], where a structured system engineering method [16] is applied to cognitive systems. This entails characterizing the desired system in terms of functions and capabilities that it should provide, specifying requirements to achieve this functionality, and then implementing these functions in an integrated system, that is then tested and compared against the requirements. Here we specify the system functionality, with an overview in Figure 2, and then list the principal requirements on the functions of the system, followed by their implementation (numbered sequentially, R1-R5).

The principal function of the narrative companion V2.0 is to discover and mediate narrative relations (argument overlap, spatial, temporal and causal continuity) [17] between members of the group that it animates, in order to aid in the creation and maintenance of the coherence of the group. The companion system will have access to a set of favorite photographs provided by group members, with short narrative descriptions. The companion will discover relations between these descriptions and create a narrative trajectory that it will use to animate a group interaction which will allow the members to discover their shared interests and to build group coherence. We now enumerate the requirements on the system that can achieve this functionality.

A. Collecting Narrative Experience from Group Members

Requirement R1 – A group of participants should be identified to participate in an interaction with the system. Prior to the interaction the group participants should provide a set of favorite photographs, and respond to a set of questions in order to provide narrative descriptions.

B. Structured organization of narrative elements

Requirement R2 – the system should have access to a structured representation of the responses to the different questions from each subject for each of their photographs. This should allow for easy access to the data according to which question was being asked, which subject was responding, which picture was being referenced.

C. Narrative Semantic Similarity Analysis

Requirement R3 - The system should perform an analysis of the narrative responses to each question and calculate the similarity between the responses to each of the photographs for all group members. For example, if Peter has a photograph that was taken on Mt. Koyasan in Japan, and Takahisa has a photograph taken at Disneyland in Tokyo, then the system should recognize a semantic similarity in the context of a question concerning where these photographs were taken.

These pairwise comparisons between the text responses should be performed to generate a complete matrix of semantic similarity (Semantic Similarity Matrix) across responses to all questions for pictures for all subjects.
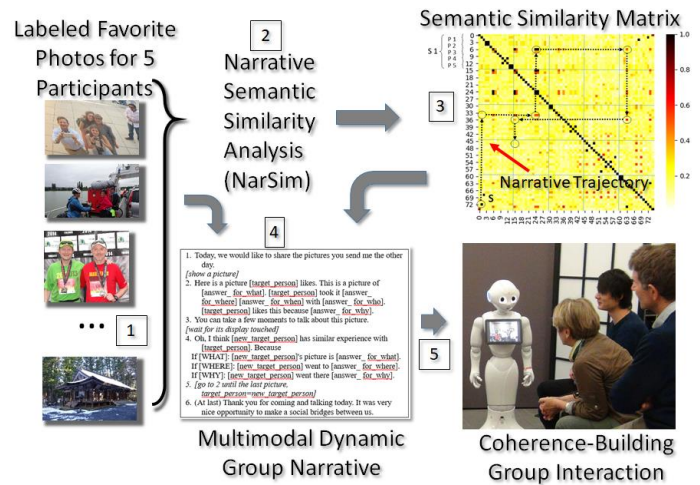

Figure 2. System overview. 1. Participants provide 5 favorite photos with responses to five questions about what, where, when, who, why. 2. These responses are compared by the Narrative Semantic Similarity Analysis (NarSim) to generate the Semantic Similarity Matrix 3. The NarSim then generates a Multimodal Dynamic Group Narrative 4 based on semantic similarities between participants' photos. 5. This is used by Pepper to animate the Coherence-Building Group Interaction with the participant group.

D. Narrative Trajectory Generation

Requirement R4 – Using the Semantic Similarity Matrix, the system should identify a trajectory that links a succession of photographs with high semantic relations. This narrative trajectory should be used by the system in a group interaction where the successive photographs are displayed with the corresponding answers to questions read out loud by the Pepper.

The trajectory should start with a photograph (e.g. Peter's photograph of Mt. Koyasan), the response to a question (e.g. from Peter. Where was this photograph taken),

then the response to the same question by another participant for one of their photos that had a high semantic similarity with the first response (e.g. Takahisa's response to a Where question that includes mention of Japan), and then the corresponding photograph (e.g. from Takahisa). This process should then repeat, i.e. identification of a response regarding this photo (e.g. from Takahisa) that has a high semantic similarity with a response from another participant, then that person's photograph, etc. This linking of people's related photos by their responses with high semantic similarity should result in a Narrative Trajectory.

E. Coherence-Building Group Interaction

Requirement R5 – Using the Narrative Trajectory, the system should animate a group interaction where group members have the opportunity to discover new links with their colleagues that have been identified by the NarSim system. At each step, the system should present the next photo, read the response to that photo, then the semantically similar response to the next photo, and then show the next photo. For each photo, the system should provide the participants adequate time to share their memories and build up group coherence before proceeding to the next response or next photo.

III. METHODS B: SYSTEM SPECIFICATION

The above requirements were used to implement the V2.0 Narrative Companion on the Pepper humanoid robot, extending the V1.0 Narrative Companion we previously developed [5].

A. Robot Platform and Native Software Environment

Pepper (illustrated in Figure 1) is a 1.2m bimanual humanoid with articulated head, arms, hands, and torso mounted on a mobile base, by Softbank Robotics, designed for social interaction with humans. The robot has 20 degrees of freedom, 4 microphones, a number of sensors (Gyro, sonar, laser, touch, bumper, 2 RGB cameras, 1 depth sensor). The torso is equipped with a 10.1 in touch display. The system communicates via Wi-Fi: IEEE 802.11, and Ethernet.

The native operating system is NAOqi OS, a Linux-based operating system that has an extensive API and significant development history extended from the NAO robot from Aldebaran. A rapid prototyping environment is provided via Choregraphe, which allows access to the NAOqi API and the construction of state-based graphical specification of complex behaviors that include dialog management, control of the tactile tablet, interaction with social media, etc.

B. Photograph Description

In order to address requirement R1 for collecting narrative experience from group members, we assembled a group of 6 participants who work together in a multidisciplinary academic research team. In order to provide the system with access to interesting elements of their past history, each person in the group was asked to provide 5 favorite photographs, and for each to respond to five questions in a structured response template. Five of the

6 participants provided photographs. Here we provide the questions and sample answers from one subject for one photo.

Q1. What is this photo? This is a photo of our summer vacation bicycle trip.

Q2. Where did you take this photo? On a ferry boat on the Seine river

Q3. When did you take this photo? Last summer

Q4. Who was with you when you took this photo? I was with my wife Jocelyne, and two daughters, Emilie and Alexia

Q5. Why do you like this photo? I have always wanted to go bicycle touring with my family, and this is a picture of that.

This satisfies the requirement R1 for collecting narrative experience from group members.

C. Distributed Autobiographical Memory

The Distributed Autobiographical Memory (DABM) represents these photographic and narrative memories of group members. The DABM is developed as a structured CSV file. Group members provided photographs and responses to questions about those photographs in a structured response template. A semi-automatic procedure was developed that reads these responses and generates entries in the DABM. The DABM can grow over time as group members add new memories to the system. In particular, based on interactions that identify new themes to be developed in the group. This satisfies requirement R2 for structured organization of narrative elements.

D. Narrative Semantic Similarity Analysis

In the DABM database, each of 5 participants provides 5 photographs and responds to the 5 questions thus yielding 125 short narrative descriptions, such as "I have always wanted to go bicycle touring with my family, and this is a picture of that." The NarSim creates a semantic representation of each response, and then for each participant, systematically compare all responses across all the other participants. The objective is to identify semantic similarity between the pictures of different participants.

In order to generate the semantic representations for each response, we exploit word2vec, a powerful method for generating multidimensional representation of words based on neighborhood relations in large corpora [9]. We employ a pre-trained version of word2vec that has been trained on a corpus of over 250K words from the 2019 Wikipedia corpus[1]. For each question response, for each word, we calculate the word vector using the Wikipedia trained word2vec model, and aggregate these word vectors into a single average vector for the response. We then compare these response average vectors to identify responses that are semantically similar. Admittedly, this method - comparing similarity of picture descriptions using the average of the vector representations of the words in the descriptions - loses the word order information. However, it preserves sufficient semantic content to provide a baseline for the SemEval Semantic Textual Similarity evaluation task [10], and is thus sufficient for our purposes.

---

[1] http://vectors.nlpl.eu/explore/embeddings/en/models/

We calculate the similarity between two response vectors as the cosine of the angle between them which varies between 1 (identical) and 0 (orthogonal). Similarity is systematically calculated for all responses to generate the Similarity Matrix. The resulting Semantic Similarity Matrix is illustrated in Figure 3. We observe a diversity of dark cells (cosines approaching 1) indicating high semantic similarity between different responses to the picture descriptions. This satisfies R3 for narrative semantic similarity analysis.

E. Narrative Trajectory Generation

The objective now is to develop an algorithm that traverses this matrix in order to generate a trajectory of photographs and descriptions that make meaningful links between the different participants. We thus search for the highest similarity value between two descriptions coming from photographs from two different participants, insert these into the narrative, and then iterate through this procedure with the constraints of not repeating the same photograph, and always making successive photographs from different people (see Table 1 for pseudo code).

```
Select the highest value in SimMatrix(i,j)
Repeat until desired narrative length
{
Identify responses i and j, and
    corresponding photographs for current
    highest SimMatrix(i,j)
Verify that the two photographs are from
    different participants
Add response i and corresponding photo,
    and response j and corresponding photo
    to the Narrative
From this cell in SimMatrix, search along
    the same row or column for the next
    highest SimMatrix(i,j) value.
}
```

Table 1. Pseudo-code for calculation Narrative Trajectory. See example in Figure 3.

In the construction of the narrative trajectory we insert the first photograph, the description of that photograph, the description of the second photograph that has high similarity, and then the second photograph. The template for creating a narrative around these descriptions is presented in Table 2. Inserting the descriptions into the template thus creates a narrative unit, with a photo and description from one participant, followed by a similar description and photo of a second participant. This procedure is then followed, by searching the Similarity Matrix for a description (and the corresponding photograph from a different participant) that is similar to this second description and adding these in an iterative manner to the narrative trajectory until the desired length has been reached.

In the initial analysis of the results of the NarSim algorithm, we detected that a number of high similarity values were detected for semantic similarity related to the two questions concerning when and with whom the picture was taken. *When* questions resulting in answers like "last summer" or "2016" which generated high scores for semantic similarity that did not reflect true similarity of the pictures.

Likewise, responses to the *with whom* questions often contained the non-specific response "with friends" which generated high semantic similarity scores that were not pertinent. We thus excluded responses to these when and with whom question, resulting in a Semantic Similarity Matrix of dimension 75 x 75, and used this to generate the narrative trajectory based on semantic similarity. An example of the resulting narrative trajectory is illustrated in Figure 3. This satisfies R4 for narrative trajectory generation.
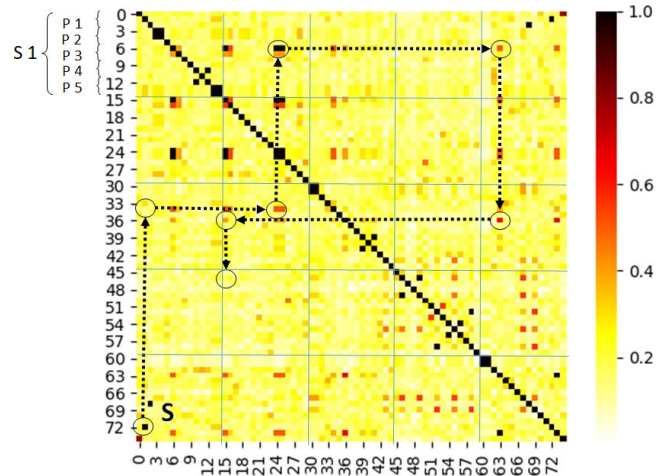

Figure 3. Semantic Similarity Matrix and example Narrative Trajectory. The 75 rows and columns correspond to the 3 responses to Q1, Q2 and Q5 for each of the 5 pictures for each of the 5 participants. Each cell is the cosine (similarity) score for the two questions. Red-dark cells correspond to high similarity scores (see scale on right). The Narrative Trajectory starts at the cell labeled S is the lower left corner of the matrix. Following the described algorithm a sequence of high similarity scores for responses to pictures from different participants are successively identified, generating the trajectory indicated by dotted lines.

1. Today, we would like to share the pictures you send me the other day.
[show a picture]
2. Here is a picture [target_person] likes. This is a picture of [answer_for_what]. [target_person] took it [answer_for_where] [answer_for_when] with [answer_for_who]. [target_person] likes this because [answer_for_why].
3. You can take a few moments to talk about this picture.
[wait for its display touched]
4. Oh, I think [new_target_person] has similar experience with [target_person]. Because
  If [WHAT]: [target_person]'s picture is [answer_for_what]and [new_target_person]'s picture is [answer_for_what].
  If [WHERE]: [target_person] went to [answer_for_where] and [new_target_person] went to [answer_for_where].
  If [WHY]: [target_person] went there [answer_for_why] and [new_target_person] went there [answer_for_why].
5. [go to 2 until the last picture,
    target_person=new_target_person]
6. (At last) Thank you for coming and talking today. It was very nice opportunity to make a social bridges between us.

Table 2. Template for generating the spoken interaction of the Narrative Companion based on the Narrative Trajectory.

F. Narrative Trajectory Group Dynamics

The final and perhaps most important component is the implementation on the Pepper of the dynamic human-robot interaction mechanism that animates the group interaction based on the Narrative Trajectory. This requires that the Pepper present the photographs and question responses in a structured manner that highlights the relations between photographs and participants.

Based on the generated Narrative Trajectory, we use the Narrative Template illustrated in Table 2 to generate a multimodal dynamic group narrative that is presented by the Pepper, which will be illustrated by a user case study. During the interactions, after the display of the next picture on the frontal screen of the Pepper, the system invites the group to discuss. A participant then informs the system to proceed in the narrative by touching the display. This satisfies R5 for coherence-building group interaction.

IV. RESULTS: SYSTEM VALIDATION AND USER CASE STUDY

We now demonstrate how the implemented system meets the initial requirements by reporting results of the system functionality in a user case study. An important aspect of systems such as NarSim is that they should function not only in theory, but in practice with participants who are not familiar with robot technology. Thus, as stated above, we recruited 6 participants from an a academic research group in our local university setting. Five provided pictures and responses as described above, and all 6 participated in the group interaction. Four of the participants were unfamiliar with the system and two were familiar with the system.

A. Coherence-Building Group Interaction Preparation

During the week prior to the interaction 5 of the 6 participants provided 5 photographs and short responses to the questions presented in paragraph II.A. On the day of the interaction the 6 participants responded to questions related to their perception of the NarSim system before and after the interaction. The questions were presented together, but correspond to three categories and are presented in Table 3.

A: General about feelings towards the robot
A1. The robot knows something about me ?
A2. The robot can understand me?
A3. The robot can help me to remember?
A4. I would like to have this robot in my house?
A5. Other members of my family would like to interact with this robot ?
B: Judgement of the robot's understanding of the photos
B1. The robot can understand the pictures that I gave it?
B2. The robot can make links between my pictures and those of the other participants?
B3. I would like to have the robot have access to all my photos?
C: Group cohesion and knowledge
C1. I feel that I know what my colleagues like to do for fun and recreation?
C2. I have a good sense of what kind of places my colleagues like to visit?
C3. I would be able to think of something fun that we could all do together with my colleagues?

Table 3. Questions asked to participants before and after the shard narrative experience. For each, participants were asked to give a score between 0 (not at all) to 5 (strongly agree).

Finally, subjects were asked to respond to the Inclusion of Other in Self survey [18] which provides an indication of the feeling of social proximity towards another.

B. Similarity analysis and Narrative Trajectory Generation

Responses to the questions for each picture concerning what, where, when, with who, and why they like the photo were analyzed by the NarSim system to generate the Similarity Matrix. Initial inspection of the matrix revealed that responses to when and who questions gave high similarity to pictures that were taken in the same year, or with "friends", which was too unspecific, so we regenerated the Similarity Matrix with responses only to what, where, why responses, illustrated in Figure 3.

Recall that the Similarity Matrix uses the learned distributed semantics from a 250K word Wikipedia corpus to discover semantic relations between the picture descriptions. These relations are then used to generate the Narrative Trajectory which follows the highest semantic links between the pictures of the different participants. An example of two descriptions that were detected to be similar and the corresponding photos is presented in Table 4, which shows how photographs taken in Thailand and Mayamar are related by the system.

```
Picture: Thailand, Person: Clement
- Question: what, Clement: Two boats on the
    shore of an island in Thailand
    Semantic Similarity 0.661778
- Question: what, Nicolas: the team working
    with in Myanmar
Picture: Myanmar, Person: Nicolas
```

Table 4. Narrative segment linking two photographs based on high semantic similarity between the answers to the *what* questions.

C. Group Dynamic Interaction with Narrative Trajectory

The crucial test of the NarSim system is the actual group dynamic interaction animated by the Pepper. In the case study we describe here, the NarSim system generated a trajectory of 9 photographs that included at least one of the photos provided by each participant. During the interaction, the participants were seated comfortably around the Pepper as illustrated in Figure 4.

Figure 4. Six participants around Pepper who is animating the coherence-building group interaction.

This interaction study was the first time that the group had ever been assembled to interact with the Narrative Companion. An excerpt of the automatically generated interaction narrative for the coherence-building group interaction is presented in the Appendix, below. The actual interaction had a total duration of over 47 minutes. Participants were highly engaged during the entire duration. The average interaction period per picture was 5 minutes 17 seconds. The individual times over the course of the 9 photographs are illustrated in Figure 5. It is interesting to observe that the duration of discussion over the successive photographs displays a tendency to increase over the course of the interaction. A video of an example part of the interaction can be seen at https://youtu.be/04fbPfsOp_8 . The interaction group was made up of native speakers of English, French, and Japanese. The common language was English, and during the interaction French was also spontaneously used. This illustrates the fluidity of the interaction.
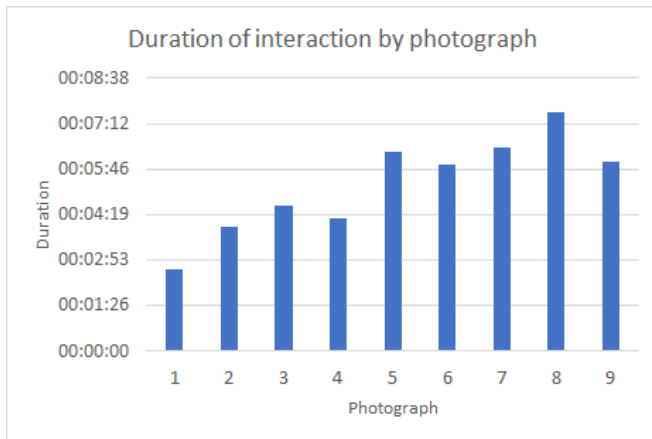

Figure 5. Duration in minutes and seconds of the discussion for each successive photograph over the unfolding of the shared narrative experience.

The extended duration of the interaction suggests that the participants were actively engaged, but this does not directly inform us concerning the effects of the interaction on the group. For this we must examine the participants' evaluation of the experience. Figure 6 illustrates the participants' responses to the questions in Table 3 before and after the coherence-building group interaction. There we see that for all three categories of question: (A). General feelings about the robot. (B) Judgement of the robot's ability to understand the photos. (C) Perception of the group cohesion and knowledge of the group, and Inclusion of Other in Self scale, there is an increase in the participants' perception of each category after the interaction. These observations are confirmed by pairwise t tests (see Figure 6 legend).

As two of the participant were familiar with the methodology of this system, we also performed a paired t-test pooling data from A-D for the four naïve subjects and verified that there were significant difference between before

and after ($p < 0.01$).

In summary, the results of the system validation and user case study demonstrate that the Narrative Companion V2.0, using the NarSim system, was able to generate an engaging group narrative that held the participant's engagement and attention for over 45 minutes of interaction. As a result of this interaction, the participants displayed an increased affinity towards the robot (as revealed by question group A and IOS), and an increased judgement that the robot could understand the photos (as revealed by question group B). Importantly, after the interaction, the participants displayed an increase in their perception of the group cohesion and knowledge of the group (as revealed by question group C).

V. DISCUSSION

We set out to develop a new narrative companion that could discover relations between human group participants and could autonomously lead those participants in a group activity that would increase the coherence of the group. To address this challenge, we exploited advances in NLP that could allow the system to find similarity between descriptions that participants provided for their favorite photos. Exploiting these similarities allowed the system to create a trajectory of photos from the different participants that were linked by similar descriptions. The system used this trajectory to create a Multimodal Dynamic Group Narrative, made up of photos and their descriptions, that it used to animate the Coherence-Building Group Interaction.
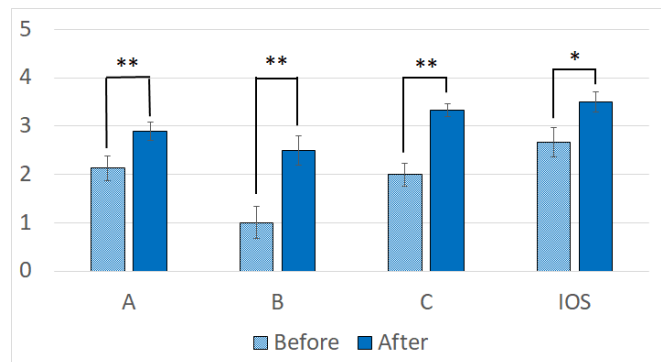

Figure 6. Results of the Questionnaire before and after the shared narrative experience. (A) General feelings about the robot. (B) Judgement of the robot's ability to understand the photos. (C) Perception of the group cohesion and knowledge of the group. (IOS) Inclusion of Other in Self. Error bars shows standard errors. * $p < 0.05$, ** $p < 0.01$.

In a validation study, this interaction lasted over 45 minutes, with participants displaying active signs of engagement and enjoyment (talking, laughing, smiling). Empirical results from a questionnaire presented before and after the interaction revealed significant increases in (a) general feelings about the robot, (b) judgement of the robot's ability to understand the photos, and, perhaps most importantly (c) perception of the group cohesion and

knowledge of the group. There was also an increase in the feeling of closeness towards the robot. Future research can work to address these dimensions – particularly the group coherence in more detail, with improved questionnaires. These results indicate that semantic knowledge from NLP can be exploited in such human robot interactions. We believe that this use of narrative companion robots to help humans get closer to each other, to help in increase group knowledge and cohesion, is an important domain where narrative robots can be of value to humans.

It is important to appreciate the semantic power available by using the word2vec model that has been trained on the Wikipedia corpus of over 250K words organized in sentences based on those that we read in Wikipedia. Wikipedia is a reasonable approximation of a general world knowledge, albeit from a particular perspective [19]. The word2vec model that is trained on this corpus thus incorporates a certain aspect of this world knowledge – particularly that of word proximity which is a good proxy for semantics. Despite this knowledge, it is possible that the NarSim system identifies narrative relations that are not the most obvious to the humans (for example, noting similarity between two descriptions that both talk about being on a river). Interestingly such cases can actually be an advantage, as we know that in human-robot dialog systems, a robot with some degree of disagreement with the human is considered to have its own intentions, and thus to be more interesting to interact with for the human [6, 7].

The important result of this research is that recent advances in semantic interpretation of language can allow social robots to engage in rich interactions with humans. Here we exploit this emerging capability to allow a robot to use this semantic experience to find hidden links between people. This is a powerful new capability for social robots, and contributes to the methods that we can employ in order to enhance people's motivation to interact with robots and with others [7]. We must stress that multiple new methods are being developed for detecting semantic similarity [10] which can be exploited in these social robots. We do not claim to advance research in semantic similarity, but rather we hope to have shown how the use of these emerging capabilities for representing semantics can be used to give robots more heart!

VI. APPENDIX:

This is an excerpt from the Coherence Building Group Interaction in the System Validation and User Case Study

Robot: (presents picture) Here is a picture Clement likes. This is a picture of An elephant next to a river. Clement took it Near Chiang Mai in Thailand Three years ago during vacation with Almost all friends. Clement likes this because one of favorite

trips. You can take a few moments to talk about this picture.

[2:36 group discussion]

Robot: Oh, I think Peter has similar experience with Clement. Because Clements picture is An elephant next to a river. And Peter went to On a ferry boat on the Seine river.

Robot: (presents picture) Here is a picture Peter likes. This is a picture of summer vacation bicycle trip. Peter took it On a ferry boat on the Seine river Last summer with wife Jocelyne and two daughters Emilie and Alexia. Peter likes this because have always wanted to go bicycle touring with family and this is a picture of that.
Please, talk about this picture.

[3:55 group discussion]

Robot: Oh, I think Nicolas has similar experience with Peter. Because Peter went to On a ferry boat on the Seine river. And Nicolas picture is a landscape of an island.

Robot: (presents picture) Here is a picture Nicolas likes. This is a picture of a landscape of an island. Nicolas took it in Japan on Zamami island in Okinawa in August 2015 during a short trip to Okinawa before going back to France with some friends. Nicolas likes this because the landscape is beautiful and the great time had in Japan during internship.
Do you have any comments for this picture?

[4:36 group discussion]

Robot: Oh, I think Takahisa has similar experience with Nicolas. Because Nicolas went to in Japan on Zamami island in Okinawa. And Takahisa went to In Disney Land in Japan.

Robot: (presents picture) Here is a picture Takahisa likes. This is a picture of a castle in Disney Land in Japan. Takahisa took it In Disney Land in Japan In Halloween season in last year with with friends. Takahisa likes this because The monument was special for Halloween and cute.
You can take a few moments to talk about this picture.

[4:12 group discussion]

REFERENCES

[1]   R. Fivush, "Constructing narrative, emotion, and self in parent-child conversations about the past," *The remembering self: Construction and accuracy in the self-narrative,* pp. 136-157, 1994.

[2]   K. Nelson and R. Fivush, "The emergence of autobiographical memory: a social cultural developmental theory," *Psychological review,* vol. 111, no. 2, p. 486, 2004.

[3]   P. Rochat, *Others in mind: Social origins of self-consciousness.* Cambridge University Press, 2009.

[4]   G. A. Fine and U. Corte, "Group pleasures: Collaborative commitments, shared narrative, and the sociology of fun," *Sociological Theory,* vol. 35, no. 1, pp. 64-86, 2017.

[5]   P. F. Dominey, V. Paleologue, A. K. Pandey, and J. Ventre-Dominey, "Improving Quality of Life with a Narrative Companion," presented at the IEEE Ro-MAN, Lisbon, Portugal, 2017.

[6]   T. Uchida, T. Minato, and H. Ishiguro, "A values-based dialogue strategy to build motivation for conversation with autonomous conversational robots," in *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2016: IEEE, pp. 206-211.

[7]   T. Uchida, T. Minato, and H. Ishiguro, "Does a Conversational Robot Need to Have its own Values? A Study of Dialogue Strategy to Enhance People's Motivation to Use Autonomous Conversational Robots," in *Proceedings of the Fourth International Conference on Human Agent Interaction*, 2016, pp. 187-192.

[8]   A.-L. Mealier, G. Pointeau, S. Mirliaz, K. Ogawa, M. Finlayson, and P. F. Dominey, "Narrative Constructions for the Organization of Self Experience: Proof of Concept via Embodied Robotics " *Frontiers in Psychology: Language,* 2017.

[9]   T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Advances in neural information processing systems*, 2013, pp. 3111-3119.

[10]   D. Cer, M. Diab, E. Agirre, I. Lopez-Gazpio, and L. Specia, "Semeval-2017 task 1: Semantic textual similarity-multilingual and cross-lingual focused evaluation," *arXiv preprint arXiv:1708.00055,* 2017.

[11]   M. Dehghani *et al.*, "Decoding the neural representation of story meanings across languages," *Human brain mapping,* vol. 38, no. 12, pp. 6096-6106, 2017.

[12]   J. Kory and C. Breazeal, "Storytelling with robots: Learning companions for preschool children's language development," in *The 23rd IEEE international symposium on robot and human interactive communication*, 2014: IEEE, pp. 643-648.

[13]   E. Barrett *et al.*, "Evaluation of a companion robot for individuals with dementia: Quantitative findings of the MARIO project in an irish residential care setting," *Journal of Gerontological Nursing,* vol. 45, no. 7, pp. 36-45, 2019.

[14]   M. Law *et al.*, "Developing assistive robots for people with mild cognitive impairment and mild dementia: a qualitative study with older adults and experts in aged care," *BMJ open,* vol. 9, no. 9, p. e031937, 2019.

[15]   P. Dominey and F. Warneken, "The basis of shared intentions in human and robot cognition," *New Ideas in Psychology,* vol. 29 no. 3, p. 14, 2011.

[16]   *NASA Systems Engineering Handbook.* NASA/SP-2007-6105 rev 1, 2007.

[17]   R. A. Zwaan, J. P. Magliano, and A. C. Graesser, "Dimensions of situation model construction in narrative comprehension," *Journal of experimental psychology: Learning, memory, and cognition,* vol. 21, no. 2, p. 386, 1995.

[18]   A. Aron, E. N. Aron, and D. Smollan, "Inclusion of Other in the Self Scale and the structure of interpersonal closeness," *Journal of personality and social psychology,* vol. 63, no. 4, p. 596, 1992.

[19]   E. Gabrilovich and S. Markovitch, "Computing semantic relatedness using wikipedia-based explicit semantic analysis," in *IJcAI*, 2007, vol. 7, pp. 1606-1611.