Talking Topically to Artificial Dialog Partners – Emulating Humanlike Topic Awareness in a Virtual Agent

Alexa Breuing and Ipke Wachsmuth

Artificial Intelligence Group
Bielefeld University, Bielefeld, Germany
{abreuing, ipke}@techfak.uni-bielefeld.de

Abstract. During dialog, humans are able to track ongoing topics, to detect topical shifts, to refer to topics via labels, and to decide on the appropriateness of potential dialog topics. As a result, they interactionally produce coherent sequences of spoken utterances assigning a thematic structure to the whole conversation. Accordingly, an artificial agent that is intended to engage in natural and sophisticated human-agent dialogs should be endowed with similar conversational abilities. This paper presents how to enable topically coherent conversations between humans and interactive systems by emulating humanlike topic awareness in the virtual agent Max. Therefore, we firstly realized automatic topic detection and tracking on the basis of contextual knowledge provided by Wikipedia and secondly adapted the agent's conversational behavior by means of the gained topic information. As a result, we contribute to improve human-agent dialogs by enabling topical talk between human and artificial interlocutors. This paper is a revised and extended version of [3].

Keywords: Automatic Topic Awareness; Embodied Conversational Agents; Human-Agent Interaction; Topic Detection and Tracking; Wikipedia

1   MOTIVATION

Topic awareness plays an important role in human conversations. Besides resolving linguistic references and ambiguities which often arise in natural language talks, it enables the interlocutors to interactionally produce coherent sequences of spoken utterances. More precisely, every spoken contribution may raise new potential topics whose actual realization depends on the co-participant's acceptance by picking up one of these topics within his or her reply [23]. Hence, a topic can be described as a *joint project* [8] as it is jointly established during ongoing conversations. Furthermore, being aware of topics helps us to touch the right subject according to the social circumstances enclosing the interactional situation. Assuming an everyday small talk conversation, for example, so-called unsafe topics such as religion and death, should be avoided [9]. Altogether, the competence to talk topically constitutes a basic requirement to carry on meaningful, flexible, and appropriate conversations with other persons.

Embodied conversational agents (ECAs) are virtual characters possessing humanlike conversational behaviors to establish an intuitive human-machine interface [6].

That is, they are capable of holding face-to-face conversations with humans by understanding and producing speech, gestures, and facial expressions. Nevertheless, they often fail to converse in great depth and hence to mutually establish a topical talk with their human opponent. In addition, many ECAs lack in simulating a sense for the adequacy of certain topics during dialog. To remedy these weaknesses, the artificial interlocutor needs to be aware of ongoing and potential conversational topics like humans.

To provide conversational agents with artificial, humanlike topic awareness in everyday interactions two main tasks need to be automatized: First, the detection of topics raised in ongoing natural language dialogs and second, the adequate integration of the resulting topic information into the agent's underlying system architecture. This paper introduces an approach tackling both tasks: We show how to connect well-established linguistic information retrieval methods with benefits originated from collaborative work provided by Wikipedia to automatically detect dialog topics. Additionally, we present how to utilize the obtained information to improve the conversational abilities of virtual computer characters regarding topic handling.

The rest of the paper is organized as follows. In the next section we introduce our notion of dialog topics establishing the basis for the present work. Subsequently, the several processes of our automatic topic detection approach are described in Section 3. Thereby we especially emphasize the application of collaborative knowledge provided by Wikipedia. Section 4 highlights the embedding of the resulting topic information into the existing architecture of the conversational agent Max. As a result, we contribute in emulating humanlike topic awareness in artificial agents as described by means of our dialog scenario in Section 5. Moreover, we present how to evaluate our model in the near future. In Section 6 we give an overview of related work before closing the paper with a short conclusion and discussion.

2   INTRODUCING DIALOG TOPICS

Assuming dialogs to be face-to-face conversations between two partners, a dialog topic emerges from a *joint activity* performed by both interlocutors [14]. That is, considering single utterances to specify a dialog topic is insufficient as they do not have topics in isolation. They rather provide topic suggestions [20]. However, the topic formulation of the particular topic is done at different levels of abstraction and from different subjective positions [23]. Speaker A, for example, might categorize a dialog about Whiskey and Brandy using the term "Alcohol", whereas speaker B might choose the term "Drinks" or "Spirits" referencing the same topic. According to this, we define a dialog topic to be an *independent, self-selected category superordinate to a co-constructed sequence of dialog contributions* [4].

2.1   Topic Shifts

A dialog topic subordinates a sequence of coherent dialog contributions as wholes [5, 23]. Hence, they generalize the concepts mentioned in these contributions to a certain degree. A potential *topic shift* in dialogs occurs, once previous concepts and concepts coming up subsequently cannot be generalized to one topic anymore. If attending to the

new concepts opens a completely different dialog topic and comes along with a drop of the present one, we refer to this kind of shift as *topic leap* [23].

On the other hand, a topic shift might happen gradually. Imagine the following dialog sequence:

A: *"In which city do you live?"*
B: *"Munich."*
A: *"Ah, then you are a fan of Bayern Munich?"*
B: *"Actually no. I like Arsenal."*

By mentioning the concept "city", speaker A suggests to talk about places. Speaker B agrees to this topic by replying with an utterance containing the concept "Munich" specifying a German city. "Munich" in turn is unrelated to the upcoming topic "Sports", however, it is conceptually closely connected to Munich's local soccer club "Bayern Munich". Thus, the dialog merges seamlessly from the topic "Places" to the topic "Sports". Hobbs calls this phenomenon *topic drift* [11].

2.2 Selection of Dialog Topics

Raising an issue requires choosing a dialog topic first. Thereby, the amount of possible topics is constrained due to the given dialog scenario, the personal relation between the dialog partners, and their cultural background. Accordingly, not every dialog topic is appropriate for everyday small talk conversations.

Referring to Schneider [21], there are three groups of basic options for topic selection:

(1) The **immediate situation** involves all topics addressing the *frame* of the dialog situation.
(2) The **external situation** represents the larger context of the immediate situation and hence of its topics.
(3) The **communication situation** refers to the conversation partners and holds private topics such as hobbies or family.

A typical small talk starts with a topic related to the immediate situation and continues with topics from the external or communication situation. Due to these social conventions, most small talk structures are very similar and ease striking up a conversation with other, especially unknown persons.

3   AUTOMATIC DETECTION OF DIALOG TOPICS

Constituting a matter of course for humans, the automatic detection of dialog topics poses a great challenge. Given a dialog situation as defined before, it has to meet several requirements. First of all, the underlying processes have to work *online*. As dialogs are continuous and demand adaptive moment-by-moment decisions, it is necessary to incessantly provide the system with information about the current topic situation directly influencing the agent's conversational behavior. Additionally, this information

has to be processed within a short time frame to guarantee humanlike reaction time. Moreover, the wide range of possible topics, for example being discussed in everyday conversations, calls for a *dynamic* handling of previously unknown contributions. This in turn presumes an access to huge amounts of previously unlearned topics and how they are correlated. According to the dynamic factor and for further reasons assigned subsequently, the online encyclopedia Wikipedia proved to be the ideal knowledge source.

3.1   Topics Provided by Wikipedia

According to our definition, dialog topics are considered to be *categories* subordinating a sequence of dialog contributions. The Wikipedia category system is composed of categories subordinating articles presented by natural language texts. Utilizing the similarity between utterance-topic relations in dialogs and article-category links in Wikipedia constitutes the basis for our dynamic topic detection approach. Generally speaking, we identify a dialog topic by mapping the several utterances to Wikipedia articles and specifying their shared Wikipedia categories as potential topics. Thus, the detection process is capable of identifying a topic t without having a priori knowledge of the domain underlying t.

   A big advantage of accessing Wikipedia for our purpose is the fact that its encyclopedic knowledge is constructed collaboratively by numerous volunteers. Hence, Wikipedia provides huge amounts of information whose maintenance is done by others. Furthermore, the resulting description and categorization of concepts reflect the participants' perception of conceptual structures and delivers insights into the human understanding of topics and their relations.

3.2   Online Detection

Within our approach, realizing an automatic topic detection mainly involves the implementation of automatic processes that identify potential topics, track ongoing topics, detect topical shifts, and label the coherent dialog sequences. To ensure an online working topic detection the first two tasks need to be performed continuously, that is on every incoming utterance. Their outcomes simultaneously affect the remaining processes. In the following, the several tasks are described in more detail. Additionally, Figure 1 gives an overview of the presented topic detection approach and illustrates the relations between its associated processes.

Identification of Potential Topics Referring to Schank (1977), an utterance said in response to an input provides both a conceptual intersection to the present dialog topic and a new conceptualization introducing potential new topics. Accordingly, to automatically identify potential topic directions, at first every single dialog contribution has to be conceptualized by identifying its contained *concept terms*. Therefore, the system first preprocesses the present utterance by means of the *Stanford Part-Of-Speech Tagger* [24]. Afterwards, all identified nouns and proper nouns are specified as concept terms. Moreover, the system extracts the verbs contained in the present utterance and transforms them to their substantive as providing potential conceptual information as well.
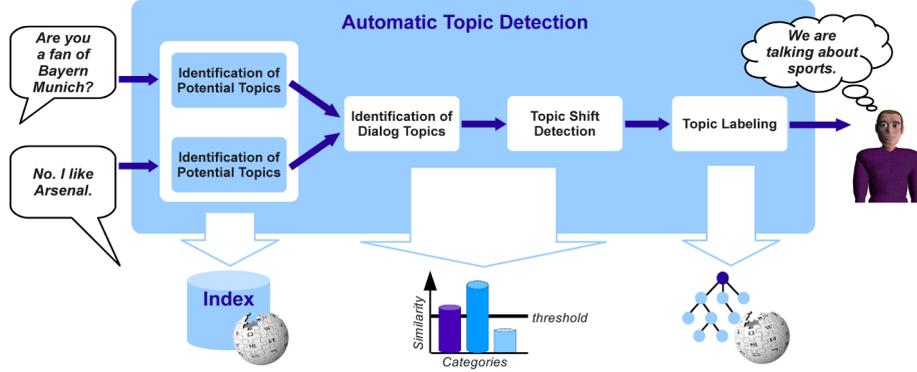
Fig. 1: Overview of the processes involved in our automatic topic detection approach.

Therefore we make use of the online dictionary *Wiktionary*. Then, the system searches for a Wikipedia article giving a concept description for the substantive. If a corresponding article can be found, as for example given for the term *"swimming"*, the substantive is considered as a concept term furthermore. In case no article is found, the substantive is not considered as a concept term as probably not providing conceptual information (like the term *"doing"*). In addition, auxiliary verbs such as *"having"* are excluded in the first place by filtering all concept terms based on a predefined stoppword list.

In order to detect named entities consisting of more than one word, adjectives and/or nouns, and proper nouns appearing successively are tested for their lexical "together-ness". Therefore we make use of the concept information provided by Wikipedia in terms of single articles [10]. More precisely, each of these potential named entities are mapped onto the set of all Wikipedia articles A_{wiki} twice: once as a whole and once noun-wise. This mapping process is accomplished via a mapping function

$$f : cterm \mapsto A_{wiki} \tag{1}$$

where *cterm* is either the potential named entity or a single noun. To realize f, we built up an *Apache Lucene* [19] search index containing documents for every Wikipedia article including information about their titles, textual descriptions, textual anchors of their incoming links, and redirects. This allows us to estimate both mappings by means of the Lucene similarity score

$$score(q,d) = \Sigma_{t \in q}(tf(t \in d) \cdot idf(t) \cdot b_f \cdot n(q,d)) \tag{2}$$

where $tf(t \in d)$ specifies the term frequency of each term $t \in cterm$ in d, $idf(t)$ indicates the general importance of t within all documents, b_f refers to the field boost in case of an exact match of *cterm* in the article title, and $n(q,d)$ combines Lucene-internal normalization factors. The outcome providing the better result determines the final composition of the concept term. By this, Wikipedia is acting as a concept identifier. As a result of the conceptualization step, a set of concept terms providing the basis for the automatic detection of potential dialog topics is determined. Thus, for the

utterance *"Ah, then you are a fan of Bayern Munich?"* the concept terms *"fan"* and *"Bayern Munich"* are specified.

One concept term can be related to more than one topic although in various extents. Within our approach, the automatic assignment of concepts to topics is implemented by mapping all concept terms to a set of predefined Wikipedia categories. Therefore, a number of categories from Wikipedia best presenting a set of topics possibly addressed in the given dialog scenario has to be specified previously. Basically, every category contained in the Wikipedia category system can be considered to present a potential dialog topic. But it is advisable to choose those categories having a high degree of abstraction as best reflecting more general topic areas such as "Sports" or "Politics".

Subsequently, for every chosen category all subordinated Wikipedia articles are extracted, that is, all articles assigned to the considered category or to at least one of its subcategories. Afterwards, the relevant information parts are stored in a second Lucene index. More precisely, documents for every predefined Wikipedia category including field specifications about its title as well as information about the titles and textual contents of their subordinated articles are set up. Thereby, articles that are related to one predefined category several times are contained accordingly often in the category document to boost its importance within the presented topic area.

To retrieve a list of categories representing possible topics sorted in descending order according to their relatedness to the concept term *cterm* we search the index for each category document *d* matching *cterm* in a query *q* on the basis of the scoring formula presented in equation 2. As a result, each concept term of the present utterance is represented as a vector within a space of predefined Wikipedia categories constituting potential dialog topics. For the rest of the paper, we refer to these vectors capturing the relative importance of the dialog topics for the considered concept term as *concept topic vectors*.

Identification of Dialog Topics As stated before, a dialog topic is established consensually from both conversation participants. That is, a single utterance does not have topics in isolation but rather provide topic suggestions [20]. Based on this idea we have to consider at least two successive utterances to define a topical intersection. Accordingly, the topic tracking process begins with the second dialog contribution.

To detect *topical overlaps* between two successive dialog contributions, we compare each of the concept topic vectors specified for one utterance with each of the concept topic vectors of the subsequent utterance separately using the *cosine similarity*. That is, we quantify the similarity between two concept terms $cterm_1$ and $cterm_2$ of successive utterances utt_1 and utt_2 on the basis of their concept topic vector representations $\overrightarrow{V}(cterm_1)$ and $\overrightarrow{V}(cterm_2)$ via

$$sim(cterm_1, cterm_2) = \frac{\overrightarrow{V}(cterm_1) \cdot \overrightarrow{V}(cterm_2)}{|\overrightarrow{V}(cterm_1)||\overrightarrow{V}(cterm_2)|} \tag{3}$$

where $cterm_1 \in utt_1$ and $cterm_2 \in utt_2$.

If the comparing process detects a significant similarity between two concept topic vectors, that is, their similarity is higher than a given similarity threshold (currently set to 0.5), a topical overlap between utt_1 and utt_2 is identified. For every topical overlap,

the involved concept topic vectors are summed up resulting in a new vector, called *dialog topic vector*. The several components in this vector provide probabilities for each predefined Wikipedia category possessing a relation to the considered concept terms. If a probability again exceeds a given probability threshold, its corresponding category constitutes the current topic of the ongoing dialog. In case the described conditions are fulfilled several times within one topic tracking process, the system is not able to determine one single Wikipedia category to be the current dialog topic but rather keeps all topic options open. Otherwise, that is if one dialog topic could be identified, the underlying dialog topic vector is included in the next identification step to keep track of this dialog topic subsequently. For this purpose, it is treated as a concept topic vector of the current utterance and is thus compared to all concept topic vectors of the following utterance to search for topical overlaps. Figure 2 graphically presents possible results of the topic tracking process for our example dialog introduced in 2.1 by means of a bar diagram. As reaching a probability ≥ 0.5 after scaling and thus exceeding the threshold represented by the horizontal line in black, the categories *"Regions"* and *"Sports"* constitute the dialog topics within this illustration.

Utterances which do not provide any concept information, like the utterance *"I know."*, have no impact on the probabilities for the several dialog topics.

Topic Shift Detection As mentioned before, we distinguish between a *topic leap* as described by Svennevig (1999) and a *topic drift* as introduced by Hobbs (1990). Based on this, systems are capable of detecting radical topic shifts enabling the particular conversational agents to generate an appropriate conversation behavior. According to this, the agent might refer to this topic leap via a suitable utterance such as *"What made you think of this topic?"*.

To distinguish between the two types of topic shift automatically, the transition from one dialog topic to the next is evaluated based on the outcomes of the topic tracking process. That is, if no topical overlap between the utterances utt_1 and utt_2 can be determined, the system detects a topic leap. In contrast, a topic drift is characterized in that topical overlaps to both the old and the new dialog topic exist during the topic transition as shown in Figure 2.

Topic Labeling To be able to refer to a dialog topic later on, for example in another dialog, a descriptive topic label has to be defined. Wikipedia provides topic labels in terms of category titles. Thus, a topic can be labeled with the title of the Wikipedia category that constitutes the current dialog topic. Thereby, the labels do not have to be mentioned during dialog before as they are already existent. However, some category titles might need to be changed to more intuitive labels. The category title "Leisure", for instance, can be replaced by "Hobbies" as the latter provides a more humanlike term for a topic raised in smalltalk conversations.
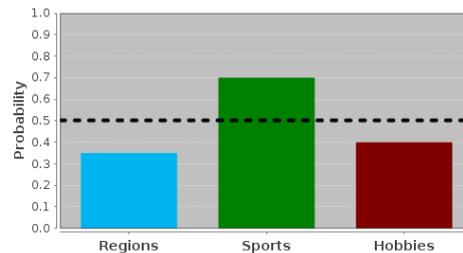
4 MAKING ARTIFICIAL AGENTS MORE TOPIC AWARE

So far, we described how to detect topics in ongoing dialog automatically by means of collaborative knowledge provided by Wikipedia. However, to emulate humanlike
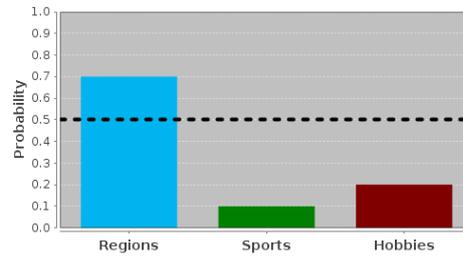
(a) Results for utterances 1 and 2.
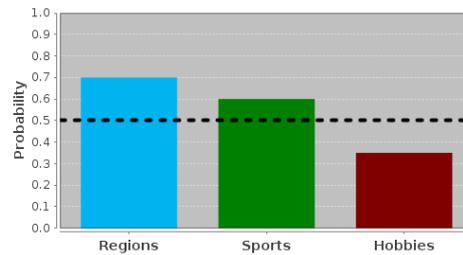
(b) Results for utterances 2 and 3.

(c) Results for utterances 3 and 4.

Fig. 2: Bar diagrams presenting results of the topic tracking process for our example dialog (see 2.1). For clarity reasons, the display adapts automatically and only shows those bars representing active topics.

topic awareness in artificial agents our presented topic detection model needs to be embedded into the agent's underlying system architecture. More precisely, the agent's conversational behavior has to be adapted by means of the gained topic information to enable coherent dialogs between human and artificial interlocutors. In this section we propose our approach for improving the conversational abilities of the ECA Max by integrating topic information into the agent's existing dialog system.

4.1   The Conversational Agent Max

Max [15] is a virtual character acting as a conversational computer interface that allows for face-to-face dialogs with humans in German language. By means of keyboard-based, textual inputs human users are able to communicate with the agent. Max is capa-

ble of responding to these inputs with spoken language realized by a synthesized voice. Figure 3 shows Max in his current state acting as a museum guide where he provides information about the exhibition and involves human visitors in everyday small talks.

Fig. 3: Max at the Heinz Nixdorf MuseumsForum in Paderborn, Germany.

4.2   Max' Existing Dialog System

The agent's verbal communication is realized by a dialog system consisting of three modules successively processing the input of the human dialog partner. In a first step the interpreter of the dialog system determines the meaning of the user's input text. The result of this analysis is then delivered to the dialog manager. By accessing the dialog knowledge, the dialog manager chooses an according answer which is sent to behavior planning afterwards. The behavior planning component translates the answer into a multimodal utterance for the virtual character.

Both the interpretation of natural language inputs and the generation of an adequate response to the user's utterance are based on a set of rules. Thereby the interpretation is composed of two steps: First, the identification of modifiers specifying the expression type such as negation, agreement, or greeting. Second, the identification of the conversational function reflecting the pragmatic and semantic meaning of the considered utterance. These processes currently employ about 1.200 rule plans which are selected and executed via pattern matching processes. These rules in turn direct the choice of an adequate response.

Due to the rule-based input interpretation covering a broad spectrum of possible utterances and an additional, Wikipedia-based question answering component [25], the agent's system never fails in computing an appropriate reply. Hence, Max never stays speechless even if an input cannot be decoded in detail. Nevertheless, the system has not yet been able to establish coherent sequences of dialog contributions as human-like topic awareness is not accessible for the agent. The integration of our online topic detection model into the ECA's system architecture is twofold: First, we contribute to improve human-agent conversations by enabling topical dialogs between human and ar-
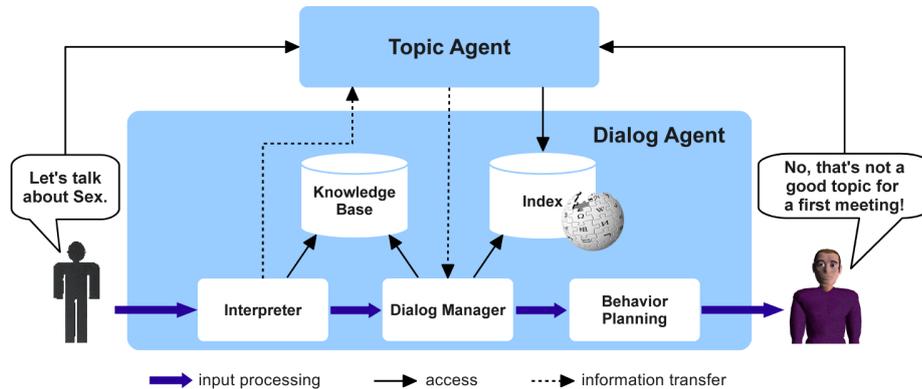
Fig. 4: Integration of our topic detection model into the existing architecture of the ECA Max.

tificial interlocutors. Second, the existing human-machine interface provides an optimal platform for the evaluation of our approach.

4.3   Integrating Topic Information

The complete system underlying the ECA Max is based on a multi-agent system composed of several interacting agents. The conversational behavior, for example, is realized via a dialog system in terms of an intelligent *dialog agent*. According to this, we built up a *topic agent* implementing the presented processes and integrated this agent into the existing system as shown in Figure 4.

The topic agent obtains every dialog contribution, that is the user's inputs as well as the agent's outputs, and constantly provides up-to-date information about the current topic situation of the ongoing dialog. It is directly connected to the dialog agent due to interdependencies. More precisely, the interpreter of the dialog agent sends its interpretation results to the topic agent which decides on the topical relevance of the considered utterance on the basis of the identified modifier. That is, if an utterance is specified as a greeting or farewell, the topic agent does not consider it as being topically relevant. Additionally, if one interlocutor proposes a dialog topic directly and the interpreter specifies a rejection in response to this suggestion, the topic agent again assigns the proposed topic to irrelevant topics. To give an example, if Max says *"Let's talk about music!"* and his human dialog partners answers with *"I don't want to talk about music!"*, the topic agent does not identify "Music" to be the dialog topic even if it was mentioned in two successive utterances.

The topic agent in turn sends the results of its topic detection process to the dialog manager which has an impact on the conversational behavior of Max. For this purpose, new dialog rules were defined allowing the agent to give information about the current dialog topic, to wonder about sudden topic shifts (i.e. topic leaps) and to restrain the search domain for the question answer component [25]. Moreover, the rules contained in the knowledge base triggering or processing topic suggestions are topically arranged

to distinguish between their adequacies according to the given dialog setting. In the following, an example extract of the resulting rule library based on the agent architecture *JAM* [13] is given.

```
/** TOPIC: REGIONS **/
Plan {
  NAME: 'askFor-homeTown'
  GOAL: get user's home town
  BODY: <act> Where are you from? </act>
  ...}
Plan {
  NAME: 'tell-favoriteCountry'
  GOAL: tell system's favorite country
  BODY: <act> I love Portugal. </act>
  ...}

/** TOPIC: SPORTS **/
Plan {
  NAME: 'reply-likingSports'
  GOAL: tell system's interest in sports
  BODY: <act> Yes, I like soccer. </act>
  ...}

/** TOPIC: POLITICS **/
Plan {
  NAME: 'askFor-politicalAttitude'
  GOAL: get user's party affiliation
  BODY: <act> What's your preferred
        political party? </act>
  ...}
```

The topical classification of the rules allows their execution based on the dialog situation. Given a first encounter, the dialog participants would not talk about their political affiliation, for instance. Accordingly, Max avoids making use of the rules dealing with so-called *unsafe topics*. That is, he neither uses such a rule pro-actively raising a topic nor reactively to answer a user question. Regarding the latter, he rather gives an evasive answer (as shown in Figure 4).

5 THE DIALOG SCENARIO

In our scenario, a human participant has a face-to-face small talk encounter with the virtual agent Max. Thereby, the human dialog partner expresses him or herself via keyboard-based text inputs whereas the artificial interlocutor answers with spoken language based on speech synthesis. Thus, the contributions of either side exist as textual information redundantizing additional speech recognition processes. Moreover, preprocessing steps to handle incomplete and non-standard sentences are not required as typed inputs mostly consist of complete sentences containing only little abbreviations and

slang expressions. However, textual inputs preclude the perception of topic ending indicators such as repetitions, pauses, laughter, etc. [12]. Thus, they can not be considered in the process of topic detection although often used in human conversation.

Enabling a coherent dialog between Max and a human user necessitates the presetting of a topical structure allowing to introduce the emulated topic awareness into the dialog and to cause a corresponding conversational behavior on the agent's part. Schneider (1988) assigns a structure to a typical small talk sequence as follows:

1. Question
2. Answer
3. Reply
4. Further turns

Furthermore, a typical small talk topically covers the immediated, external, and communication situation [21]. In their study, Endrass et al. (2011) identified a typical distribution of these topics within a dialog between Germans. Thus, Germans address less of the immediated and approximately equivalent of the external and communication situation during small talk. According to these findings, and considering the conditions arising from the fact that Max is situated in an university environment, the beginning of topical small talks with the conversational agent Max is structured as follows: In his first turn, Max asks the interlocutor for his or her subject of study as most potential dialog partners are students. Subsequently, the agent tries to find out the interlocutor's origin. If successful, Max is able to determine the interlocutor's favorite football club from this knowledge and to continue with the football topic. In case the human does not want to talk about football or sports in general, he or she has the opportunity to suggest another topic. Thus, the first dialog topics are solely initiated by Max. This is important insofar as this scenario also establishes a basis for the development of a personal memory for the agent as introduced by Mattar and Wachsmuth [18]. This requires the gathering of a lot of information concerning social categories about the human interlocutor.

5.1   Planned Evaluation

Upon successfully completing a preliminary evaluation identifying the topics of newspaper articles, which has shown an average accuracy of 61.0 [4], we plan to accomplish a more adequate evaluation considering and addressing the introduced dialog scenario. Accordingly, we searched for a corpus comprising dialog information of German small talks occurring during first encounters between two persons. The CUBE-G corpora [9] provides analyzed records of 21 first interaction scenarios each between a student and a professional actor and each lasting around five minutes. Amongst others, the dialogs were tested for the amount of topics and topic shifts which is why the corpus contains topical annotations for each recorded small talk. Thus, the CUBE-G corpus presents the optimal basis for our following evaluation.

In preparation for the planned evaluation, we already determined a list of predefined main categories that represent typical dialog topics for everyday small talks. Thereby, we omitted so-called *unsafe topics* (see Section 1) and especially focused on topics

Main Category	
Science	Economics
Family	Education
Studies	Literature
Mass media	Music
Arts	Health
Ecology	Digital media
Sports	Occupations
Fashion	Food and drink
Leisure	Transport
Intimate relationships	Regions

Table 1: List of predefined main categories adequate for our dialog scenario.

raised in the given university scenario. Table 1 shows the resulting list of main categories. Moreover, we downloaded the German database dump from May 14, 2011 and built up a Lucene index containing all information parts relevant for our purpose.

The next step is the preprocessing of the corpus in that incomplete sentences and expressions are completed to adapt the recorded utterances to the conditions given by the fact that human-sided utterances are based on keyboard inputs. Then, we will accomplish the evaluation by automatically identifying the dialog topics and topic shifts within the CUBE-G interactions by means of our proposed method to subsequently compare the results with the manual annotations included in the corpus. If showing promising performance, a user study evaluating the application of emulated human topic awareness in the agent Max' conversational behavior will be scheduled next.

6   RELATED WORK

A lot of work has been carried out on offline topic identification. A prevalent model was developed in the context of the Topic Detection and Tracking (TDT) research program [1]. Within the TDT research, Allan determined five tasks (i.e., Story Segmentation, First Story Detection, Cluster Detection, Tracking, and Story Link Detection) for detecting the several topics outlined in a text-based newscast. Further offline approaches compute the coherence between documents via similarity measures (e.g., [17, 26]). Others rank Wikipedia articles according to their relevance to a given text fragment, for example via text classification algorithms [10] or by simply exploiting the Wikipedia article titles and categories [22]. One recent approach uses the Wikipedia category network as a conceptual taxonomy and derives a directed acyclic graph for each document by mapping terms to a concept in the category network [7].

Approaches for the online identification of topics in natural language dialogs are rare. One work realizing a "Dynamic Topic Tracking" of natural language conversations between a human and a robot roughly adapted the five tasks from the TDT project (see above) to make the robot more situation aware in human-robot interaction [16]. Thereby the amount of topics and the according topic names are created dynamically by gathering the topic names from content words most occurring in the dialog utterances.

On the contrary, existing taxonomies can serve as a source for topic labels, for example derived from the online encyclopedia Wikipedia [4, 25]. Furthermore, conversation clusters visually highlight topics discussed in conversations using Explicit Semantic Analysis based on Wikipedia articles [2].

7   CONCLUSION AND FUTURE WORK

We presented an approach for the automatic emulation of humanlike topic awareness in ongoing small talk dialogs to extend the conversational abilities of a virtual agent in human-agent interactions. More precisely, we proposed solutions for both tasks the automatic identification of dialog topics and the integration of the resulting topic information into the agent's existing system architecture. The several associated processes fulfill the requirements given by a face-to-face encounter between a human and a conversational agent and enable both a coherent and socially adequate dialog between the human and the artificial interlocutors. Thereby, we exploit Wikipedia knowledge and hence the benefits originated from collaborative work (namely the existence of information whose maintenance and expansion is carried out by numerous volunteers and the reflection of the participants' common perception of conceptual structures).

In future, we will extend our approach by detecting and linking topical affiliations to previous dialog topics to handle short *side trips* to past topics. Moreover, we will resolve ambiguities by taking into account the current dialog topic to influence the concept detection process.

ACKNOWLEDGEMENTS

References

1. J. Allan. *Topic Detection and Tracking: Event-based Information Organization.* Kluwer Academic Publishers, 2002.
2. T. Bergstrom and K. Karahalios. Conversation clusters: Grouping conversation topics through human-computer dialog. In *Proceedings of the International Conference on Human Factors in Computing Systems (CHI09)*, 2009.
3. A. Breuing and I. Wachsmuth. Let's talk topically with artificial agents! providing agents with humanlike topic awareness in everyday dialog situations. In *Proceedings of the 4th International Conference on Agents and Artificial Intelligence (ICAART)*, pages 62–71, 2012.
4. A. Breuing, U. Waltinger, and I. Wachsmuth. Harvesting Wikipedia knowledge to identify topics in ongoing natural language dialogs. In *Proceedings of the 2011 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, pages 445–450. IEEE Press, 2011.

5. W. Bublitz. Topical coherence in spoken discourse. *Studia Anglica Posnaniensia*, 22:31–51, 1989.
6. J. Cassell, T. Bickmore, L. Campbell, H. Vilhjálmsson, and H. Yan. Human conversation as a system framework: Designing embodied conversational agents. In J. Cassell, J. Sullivan, and E. Churchill, editors, *Embodied Conversational Agents*, pages 29–63. MIT Press, 2000.
7. C. A. Chahine, N. Chaignaud, J.-P. Kotowicz, and J.-P. Pécuchet. Conceptual indexing of documents using Wikipedia. In *Proceedings of the 2011 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, pages 195–202, 2011.
8. H. H. Clark. *Using Language*. Cambridge Univ. Press, 1996.
9. B. Endrass, M. Rehm, and E. André. Planning small talk behavior with cultural influences for multiagent systems. *Computer Speech and Language*, 25(2):158–174, 2011.
10. E. Gabrilovich and S. Markovitch. Computing semantic relatedness using Wikipedia-based explicit semantic analysis. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI-07)*, 2007.
11. J. Hobbs. Topic drift. In B. Dorval, editor, *Conversational Organization and Its Development*, pages 3–22. Ablex Publishing, 1990.
12. M. Howe. Collaboration on topic change in conversation. *Kansas Working Papers in Linguistics*, 16:1–14, 1991.
13. M. J. Huber. JAM: A BDI-theoretic mobile agent architecture. In *Proceedings of the Third International Conference on Autonomous Agents (Agents99)*, pages 236–243, 1999.
14. D. Jurafsky and J. H. Martin. *Speech and Language Processing*. Pearson Prentice Hall, 2009.
15. S. Kopp, L. Gesellensetter, N. Krämer, and I. Wachsmuth. A conversational agent as museum guide – Design and evaluation of a real-world application. In *Proceedings of Intelligent Virtual Agents (IVA 2005)*, 2005.
16. J. F. Maas, T. Spexard, J. Fritsch, B. Wrede, and G. Sagerer. BIRON, what's the topic? A multi-modal topic tracker for improved human-robot interaction. In *Proceedings of the IEEE International Workshop on Robot and Human Interactive Communication (ROMAN)*, pages 26–32, 2006.
17. J. Makkonen, H. Ahonen-Myka, and M. Salmenkivi. Simple semantics in topic detection and tracking. *Information Retrieval*, 7(3-4):347–368, 2004.
18. N. Mattar and I. Wachsmuth. Who are you? On the acquisition of information about people for an agent that remembers. In *Proceedings of the 4th International Conference on Agents and Artificial Intelligence*, pages 98–105, 2012.
19. M. McCandless, E. Hatcher, and O. Gospodnetić. *Lucene in Action*. Manning, 2 edition, 2010.
20. R. C. Schank. Rules and topics in conversation. *Cognitive Science*, 1(4):421–441, 1977.
21. K. P. Schneider. *Small Talk: Analysing Phatic Disourse*. Hitzeroth, 1988.
22. P. Schönhofen. Identifying document topics using the Wikipedia category network. In *Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence (WI'06)*, 2006.
23. J. Svennevig. *Getting Acquainted in Conversation*. John Benjamins Publishing, 1999.
24. K. Toutanova and C. D. Manning. Enriching the knowledge sources used in a maximum entropy part-of-speech tagger. In *Proceedings of the Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora*, pages 63–70, 2000.
25. U. Waltinger, A. Breuing, and I. Wachsmuth. Interfacing virtual agents with collaborative knowledge: Open domain question answering using Wikipedia-based topic models. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI 2011)*, pages 1896–1902. AAAI Press, 2011.
26. X. Zhang and T. Wang. Topic tracking with dynamic topic model and topic-based weighting method. *Journal of Software*, 5(5):482–489, 2010.