

Harvesting Wikipedia Knowledge to Identify Topics in Ongoing Natural Language Dialogs

Alexa Breuing, Ulli Waltinger, Ipke Wachsmuth
Faculty of Technology
Artificial Intelligence Group
Bielefeld, Germany
{abreuing, uwalting, ipke}@techfak.uni-bielefeld.de

Abstract—This paper introduces a model harvesting the crowd-sourced encyclopedic knowledge provided by Wikipedia to improve the conversational abilities of an artificial agent. More precisely, we present a model for automatic topic identification in ongoing natural language dialogs. On the basis of a graph-based representation of the Wikipedia category system, our model implements six tasks essential for detecting the topical overlap of coherent dialog contributions. Thereby the identification process operates online to handle dialog streams of constantly changing topical threads in real-time. The realization of the model and its application to our conversational agent aims to improve human-agent conversations by transferring human-like topic awareness to the artificial interlocutor.

Index Terms—Wikipedia; Information Retrieval; Human-Agent Interaction; Topic Identification

I. INTRODUCTION AND MOTIVATION

As a product of collaborative work carried out by thousands of volunteering editors, Wikipedia provides a collection of encyclopedic knowledge in a magnitude never existed before. We explore the use of Wikipedia as an external source of taxonomically organized concept knowledge to improve the conversational abilities of an artificial agent. More precisely, we aim for making conversational agents more context and topic aware in natural language dialogs by connecting linguistic information retrieval methods with benefits originated from collaborative writing.

Imagine the following dialog section:

A: “*I’d like to go to the stadium tomorrow. Bayern Munich is playing. Would you like to join me?*”

B: “*No, I’m sorry, I don’t like soccer.*”

For humans, it is quite easy to assess what the dialog is about: “Sports”. Moreover, humans are able to track ongoing topics, to detect topic shifts, to refer to topics via labels, to return to already closed topics, and to reflect the history of raised topics during conversation. This topic awareness constitutes a crucial factor for the success of a meaningful conversation between human interlocutors. It enables the participants to resolve linguistic references and ambiguities which often arise in natural language dialogs. The term “stadium”, e.g., can be assigned to many different concepts. Being aware of the current topic “sports” helps to have “sports venues” in mind. Furthermore, topic awareness helps the participants to produce and exchange coherent sequences of dialog contributions [1]. These contributions in turn may raise new potential topics

whose continuation depends on the co-participant’s acceptance by not returning to the original topic [2]. Thus, topics are interactionally established and enable versatile and flexible conversations.

To achieve such topically coherent conversations in dialogs with an artificial agent, human-like topic awareness needs to be transferred to the artificial interlocutor. This challenge involves two main tasks: First, the identification of dialog topics and second, the adequate integration of the resulting topic information into the agent’s underlying system architecture. Performing the first task implies the detection, tracking, and labeling of conversational topics. During these steps additional information about topic shifts and topic detailedness based on their duration during conversation need to be detected. Finally, a history of tagged dialog sections provides information about the entire topic spectrum of the analyzed conversation. In the second task, the derived information is used to topically adapt the agent’s conversational behavior, e.g., by limiting the range of possible answers to one question [3]. Both main tasks strongly interdepend and interact. The focus of the present paper is on the first task, the identification of dialog topics.

Basic requirement for realizing automatic topic identification techniques is the access to conceptual knowledge, i.e., knowledge about concepts and how they are correlated. Wikipedia offers conceptual information in terms of natural language articles arranged in a taxonomy-like category system. This paper introduces a model realizing automatic topic identification in natural language dialogs on the basis of a graph-based representation of the Wikipedia category system.

The benefit of accessing Wikipedia as an external knowledge base is multiple: First, the model harvests existing information whose maintenance and expansion is carried out by numerous volunteers. Thus, new or changing concepts and relation structures will directly become available by simply accessing the latest Wikipedia version. Second, the textual content and categorization of the articles describing single concepts reflect the participants’ (and thus the human) common perception of conceptual structures and allows to infer deeper topic-related insights as described in this paper.

The rest of this paper is organized as follows. In the next section, a precise definition of topic in human conversation is presented. Subsequently, an overview of existing work related to topic identification is given in Section 3. Section

4 introduces the category system of Wikipedia and its benefits for our purpose. Section 5 forms the main section presenting our model. Besides introducing the basic tasks for Wikipedia-based topic identification in dialogs, the workflow of the processes defined in the model is demonstrated. In Section 6 the present implementation steps arising from the model are described as well as first evaluation steps and results. The paper ends with a conclusion and an outlook presenting ideas for future work.

II. TOPIC DEFINITION

According to Bublitz [4], a topic is not identical with individual contributions, but it is an independent, continuous category distinguishable from them and superordinate to them. Extending Bublitz's original definition [4], our new topic definition needs to be sufficient for all our model requirements. More precisely, we included the perspective of Svennevig [2] and Guotsos [5] by considering a topic to be a *self-selected* category. In contrast, the dialog section belonging to a single topic have to be *co-constructed* and thus have to enclose a *sequence* of at least two dialog contributions as stated by Schank [6] and Clark [7]. Additionally noting these considerations leads us to the following definition of topic in natural language dialogs:

A dialog topic is an independent, self-selected category superordinate to a co-constructed sequence of dialog contributions.

Thus, a topic in conversation consists of two parts: First, the dialog utterances belonging to one topic; second, the category subordinating and comprising these utterances. While the sequence of topic contributions depends on every dialog partner, the topical category is selected independently from both, the other dialog partners and the words occurring in the contributions.

III. RESEARCH IN TOPIC IDENTIFICATION

The conceptualization of a model for topic identification is strongly dependent on whether or not the topics have to be identified online. An *online* identification process is able to handle previously unknown contributions gradually accessible, e.g. occurring in ongoing conversations. Moreover, it provides adequate results subsequent to each contribution and adapts dynamically to new dialog topics. In contrast, the content to be analyzed may be available as a whole in the forefront of the identification process, e.g., in terms of a daily newspaper. In this case, the topic identification works *offline* and processes the entire content before providing identification results. Information about topic shifts, number of topics, topic durations, etc. can be factored directly in the ongoing topic identification process.

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 [8]. 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., [9], [10]). Others rank Wikipedia articles according to their relevance to a given text fragment, e.g., via text classification algorithms [11] or by simply exploiting the Wikipedia article titles and categories [12].

Approaches for online identifications of topics differ according to their application scenarios. For instance, Eagle et al. present a method for situation understanding in natural face-to-face conversations [13]. This method combines commonsense knowledge with keyword spotting and context information from mobile devices to identify fine-grained conversation topics. By contrast, conversation clusters visually highlight topics discussed in conversations using Explicit Semantic Analysis based on Wikipedia articles [14].

One approach tackling a challenge very similar to our dialog setting is applied to the BIRON robot platform [15]. It realizes a "Dynamic Topic Tracking" of natural language conversations between a human and a robot. Therefore, five tasks roughly adapted from the TDT project (see above) are accomplished to make the robot more situation aware in human-robot interaction. The amount of topics and the according topic names can be created dynamically. This dynamic topic labeling can be realized by gathering the topic names from content words most occurring in the dialog utterances [16]. On the contrary, existing taxonomies can serve as a source for topic labels, e.g., derived from the Wikipedia category system [17], [18].

In our approach, the model employs two main ideas presented in this section. First, we divide the identification process into several tasks similar to those of the TDT project. Second, we utilize Wikipedia as a basis to implement these tasks. The former requires the extension, re-sort, and adaptation of the TDT tasks according to their online application in ongoing conversations. The latter requires a machine-readable representation of the information necessary for the task executions.

IV. CATEGORY SYSTEM OF WIKIPEDIA

According to our topic definition, 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 dialog-based topic-utterance and Wikipedia-based category-article relations constitutes the main idea of our model realization. Figure 1 depicts the Wikipedia category graph composed of article nodes, category nodes, and links between them.

Consistent with our definition, the graph-based representation provides numerous topics represented by the several Wikipedia categories. Their subordinated articles in turn represent the concepts crucial for the particular topic. Thus, identifying a dialog topic for a sequence of utterances means mapping the addressed concepts to Wikipedia articles and identifying their common category. That is, the identification model is not simply considering the terms contained in the

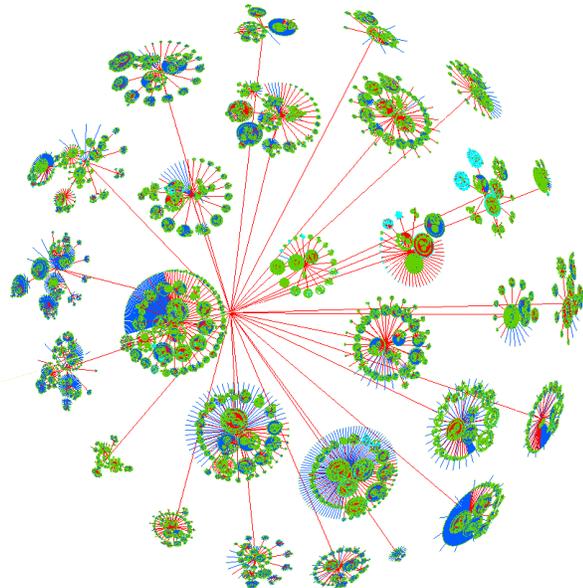


Fig. 1. A 3D graph-based representation of the Wikipedia category system providing more than 88,000 topics.

dialog utterances but additionally accesses an existing taxonomy. This enables identifying a topic t without having a priori knowledge of the domain underlying t . Moreover, a topic can be assigned with a label which wasn't mentioned explicitly during the dialog (e.g., the model is able to infer "sports" as a topic solely based on the mentioned terms "football", "basketball", "fitness", etc.).

V. TOPIC IDENTIFICATION IN HUMAN-AGENT DIALOGS

Based on the tasks made up in the TDT project [8], we propose six tasks necessary for automatically generated, human-like topic awareness in ongoing natural language dialogs. The order of the following presentation of these tasks corresponds to the order of their application to each utterance expressed in the considered dialog. Depending on the outcome of the several task executions, the task order may terminate after any particular task. More detailed information on this can be gathered from the following task specifications.

A. Task 1: Concept Detection

In accordance with Schank [6], single sentences are supposed to introduce possible concepts acting as topic suggestions. That is, the first task of our model consists in the conceptualization of dialog utterances to infer topic suggestions providing the basis for the definition of an overall dialog topic. Wikipedia provides numerous concept definitions in terms of textual articles. Thus, the *Concept Detection* task is performed by mapping *concept terms* included in the considered utterance to corresponding Wikipedia articles. Given C as the set of all concept terms of the utterance and A as the set of all Wikipedia articles, we define the mapping function

$$f_{map} : C \rightarrow A, concept \mapsto (art_1, art_2, \dots, art_n) \quad (1)$$

where $concept \in C$. and $art_i \in A$.

In our model, we define a concept term to consist of at least one noun or proper noun as well as its associated adjective. Hence, in a first step all nouns and adjectives of the current utterance are specified using the Stanford tagger [19]. Adjectives and/or nouns appearing successively are combined to one concept term initially. Subsequently, f_{map} is realized by means of an *Apache Lucene* [20] search index consisting of textual information about all Wikipedia articles. To retrieve a list of articles sorted in descending order according to their relatedness to the concept term *concept* we search the index for each article document d matching *concept* in a query q on the basis of Lucene's similarity scoring formula

$$score(q, d) = \sum_{t \in q} (tf(t \in d) \cdot idf(t) \cdot boost_f \cdot norm(q, d)) \quad (2)$$

where $tf(t \in d)$ specifies the term frequency of each term $t \in concept$ in d , $idf(t)$ indicates the general importance of t within all documents, $boost_f$ refers to the field boost in case of an exact match of *concept* in the title, redirects, or incomings of the article, and $norm(q, d)$ combines Lucene-internal normalization factors. For concept terms consisting of more than one word, the mapping operation is accomplished twice: First, the concept term is considered as a whole; second, the several words are considered separately. The outcome with the better result determines the final composition of the concept term. By this, Wikipedia is acting as a concept identifier. Subsequently, we extract the top five articles determined for one concept.

As a result of the *Concept Detection* task, the most likely concepts acting as topic suggestions are identified by means of Wikipedia articles. Corresponding to the utterance of speaker A in our example dialog, two topic suggestion "stadium" and "Bayern Munich" are identified which are mapped to the "Stadium" and to the "FC Bayern Munich" article. The results can be demonstrated graphically by highlighting these articles as shown in Figure 2.

B. Task 2: Local Topic Tracking

Like Blubitz [4] and Downing [21], we distinguish between a *local* and a *global topic*. However, we define a *global topic* to not cover the whole discourse but to cover a sequence of topically associated utterances within a dialog (equatable with our definition of a dialog topic). In contrast, we specify a topic to be *local* if it is a topic in progress. Thus, a local topic can be seen as a preliminary stage of a global topic. The intention of this task is to track the local topic to be aware of the ongoing topic situation. Applied to our approach, the topic affiliation of an utterance with the current local topic is analyzed. More precisely, this task tests whether or not the recent topic suggestions are also subordinated by the present local topic. As long as no local topic exists, a new one is defined based on the previous topic suggestions and the current ones. That is, the *Concept Detection* task has to be performed at least two times to enable the definition of a local topic.

We reduce the number of potential topics by reducing the number of categories serving as topics. Therefore, we define a manageable amount of *main categories* like "History",

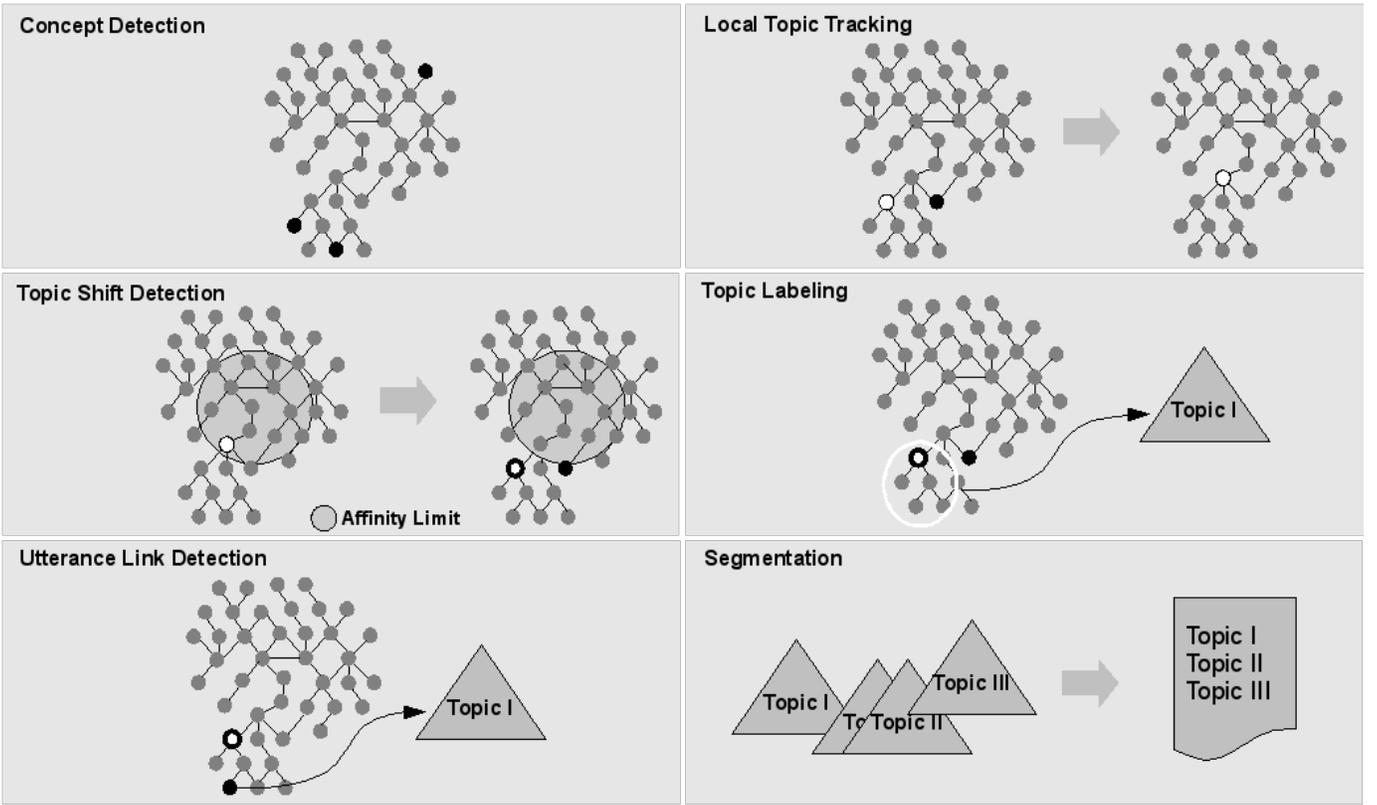


Fig. 2. The outcomes of the several tasks for automatic topic identification in dialogs exemplified by means of a simplified graph-based representation of the Wikipedia category system. The outer nodes represent articles, the remaining nodes represent categories. Thereby, black nodes highlight articles representing topic suggestions, white nodes highlight the recent local topic, and the dialog topic is also highlighted in white but additionally encircled in black.

“Music”, “Sports”, etc. By assigning all articles to main categories we can calculate the main category with the highest popularity by summing up all article weights returned by the Concept Detection task. Accordingly, we map each article art to a list of main categories in a preprocessing step via

$$f_{map} : A \rightarrow MC, art \mapsto (mCat_1, mCat_2, \dots, mCat_n) \quad (3)$$

where $art \in A$ and $mCat_i \in MC$. We realize such mapping by walking up the category graph starting from the considered article and ending in main categories. Instancing the articles resulting from the conceptualization of our example dialog, we get “Arts”, “History”, “Society”, “Technology”, and “Sports” as main categories for the article “Stadium”; “Organizations” and “Sports” as main categories for the article “FC Bayern Munich”; “Sports” and “History” as main categories for the article “Association football”. Subsequently, the basic category possessing the highest weight by summing up the Lucene scores of their associated articles is defined to be the local topic. That is, our mapping function is based on the similarity between the dialog sequence and a main category defined as

$$sim(utt, mCat) = \frac{\sum score(q, d)_{mCat}}{\sum score(q, d)_{MC}} \quad (4)$$

where utt is the utterance, $\sum score(q, d)_{mCat}$ represents the sum of the Lucene scores (weights) described in equation 2 of all articles serving as concepts of the utterance and having $mCat \in$

MC as their main category, and $\sum score(q, d)_{MC}$ represents the sum of the Lucene scores (weights) of all articles together. Due to the fact that all concepts of the example dialog has “Sports” as a main category in common, here the local topic is “Sports”. In case a further utterance arises, its topic suggestions are tested against the local topic. So if speaker A adds

A: “Oh really? And what about basketball?”

the recent local topic persists as the conceptualization gives the article “Basketball” whose main categories are “Sports” and “History”.

C. Task 3: Topic Shift Detection

In order to perceive topical shifts automatically, the first utterance belonging to a new topic needs to be detected. Hence, an *affinity limit* regarding topical closeness needs to be defined limiting the generality of a topic. Both, the definition of the affinity limit and the detection of topic shifts are tackled by the *Topic Shift Detection* task.

An affinity limit $alim$ for the category graph is realized by means of a category level presenting the top limit for local topic detection. It prevents defining the category most general cat_m (“Articles” in case of the English Wikipedia) to be the topic of the entire dialog. Hence, we define $alim \in C$. Once the previous task defines a new local topic top'_l as a category above the affinity limit, it is reset to the previous category.

That is, if *alim* is determined according to its distance to *cat_m* as illustrated in Figure 2 and if

$$\min_{top'_l, cat_m} len(top'_l, cat_m) < \min_{alim, cat_m} len(alim, cat_m) \quad (5)$$

where *len* is defined to be the distance between two nodes, then *top'_l* is reset to *top_l*. The results of the recent utterance conceptualization are set to provide topic suggestions for a new topic. Thus, an utterance like

A: “*Okay. But we could go to the theater instead!*”

would open a new topic. The previous local topic “Sports” is kept alive for the present allowing to return to this topic if the utterance actuating a new topic is a topical outlier only.

D. Task 4: Topic Labeling

A topic is not concrete until it is closed. That is, the detection of a topic shift confirms the stability of the previous topic. At this point, the recent local topic *top_l* becomes a dialog topic *top_d*. To be able to refer to this dialog topic afterwards, a descriptive topic label has to be defined. On the basis of the Wikipedia category graph, the label for the dialog topic is taken from the title of the corresponding Wikipedia category. Thereby, the labels do not have to be mentioned during dialog before as they are already existent in the graph.

E. Task 5: Utterance Link Detection

During conversation, dialog partners often refer to, or pick up, past topics. The purpose of this task is to search for topical affiliations to previous dialog topics and to link the present utterance accordingly where appropriate. A definite return to the past topic again depends on the dialog partner’s response.

Within our model, a topic resumption is identified if articles representing new topic suggestions are superordinated by a category representing an already closed topic. By this we are also able to handle short *side trips* to past topics. Regarding our example dialog, the utterance

A: “*Okay. But we could go to the theater instead!*”

can be seen as such side trip if B answers with

B: “*Oh no, then I’d prefer to go to the swimming pool.*”.

F. Task 6: Segmentation

Once a dialog is completed, its discussed topics need to be summarized, sub-divided and/or rearranged if necessary. By this, the whole dialog is divided into topically meaningful segments. For this purpose, we first need to segregate the graph sections from the category graph ranging from the dialog topic to its subjacent articles. This yields a number of partial hierarchies (trees) of topics discussed in the dialog (see Figure 2). As mentioned by Bublitz [4], these topics ideally stand beside one another as discrete blocks. On the contrary, they frequently overlap or receive later additions, etc. This relative lack of order typical for everyday conversations needs to be re-sorted. As a result, the *Segmentation* task constructs a topic history providing information about all topics discussed during dialog. Storing such topic history enables the agent to refer to topics of the past dialog at a later date, e.g., in another conversation.

VI. EVALUATION

We downloaded the German database dump (*dewiki*) from January 11, 2011 and built up a Lucene index containing certain information parts relevant for our purpose. That is, we set up documents for every Wikipedia page including field specifications about its ID, title, textual content, IDs of its superordinate categories, associated redirects, and incoming links from other Wikipedia pages. As a result, we created an index consisting of documents for about 80.000 categories and over 1,000,000 articles.

Regarding the evaluation of the topic identification and topic labeling tasks, existing corpora with topic-labeled dialogs are hardly available. We therefore compiled two different datasets, by mapping a set of natural language articles from the German online newspaper *Zeit Online*¹ to a restricted number (20) of main categories (see Table 1 for an overview of the used categories). These main categories are all Wikipedia categories which reflect the news categories provided by the newspaper while subordinating more than 50 subcategories.

We conducted two different experiments: First, we evaluated the accuracy of the topic identification on the basis of 2,000 single sentences representing individual dialog contributions. Second, we evaluated the topic identification accuracy on the basis of the first 20 article sentences representing an entire dialog sequence. Therefore, we extracted a total of 100 sentences from five different newspaper articles for each of the 20 main categories. Thereby the several topic affiliations of the sentences are already specified by means of the article segmentation provided by the online newspaper.

A. Topic Affiliation of Single Sentences

By evaluating single article sentences, we allow for considering similar conditions as given in natural language dialogs consisting of single utterances. Table 1 presents the number of correctly identified topic affiliations. The results show how variously difficult it is to identify the several topics by means of our approach. This might be associated with the given structure of the Wikipedia category graph. Another explanation might be that some topics are more related to each other which hinders their topical differentiation. Additionally, certain topics show higher frequencies for topic indicating terms than others.

B. Topic Affiliation of Articles

To evaluate the topic identification by means of the first 20 article sentences representing (longer) dialog sequences, the topic affiliations of the single sentences were counted. If the score for the correct topic was the highest, we defined the article to be topically estimated correctly. The better results (presented in Table 1) with an average accuracy of 61.0 are consistent with our assumption in that a higher number of coherent sentences can be topically identified more easily as side trips to other topics are of less significance. This again is conform to Schank’s theory [6] and to our topic definition in that single sentences do not have topics in isolation.

¹<http://www.zeit.de/>

Main Category	Sentence Accuracy	Article Accuracy
Occupations	16.0	40.0
Video game	14.0	40.0
Food and drink	57.0	100.0
Family	20.0	40.0
Finance	39.0	100.0
History	22.0	40.0
Health	45.0	100.0
Universities and colleges	29.0	80.0
International Politics	19.0	40.0
Politics of Germany	21.0	40.0
Internet	25.0	60.0
Arts	13.0	20.0
Literature	28.0	80.0
Mass media	27.0	60.0
Fashion	21.0	60.0
Music	36.0	100.0
Ecology	23.0	40.0
Schools	25.0	80.0
Sports	30.0	80.0
Companies	23.0	20.0

TABLE I
EVALUATION RESULTS.

Nevertheless, other Wikipedia-based approaches (e.g., [17], [11]) focusing on characterizing documents as a whole show better results as not considering single sentences separately.

VII. CONCLUSION AND FUTURE WORK

We presented an approach tackling two challenges in the fields of automatic topic identification: First, the transfer of the well-established TDT tasks to ongoing natural language conversations. Second, the utilization of an existing taxonomy. As a result, we obtain a model consisting of six tasks necessary for topic awareness in natural language dialogs by means of crowd-sourced encyclopedic knowledge provided by Wikipedia. Thereby we exploit the benefits provided by these components: We make use of the breakdown in separate tasks enabling a gradual implementation and evaluation. Additionally, we make use of the Wikipedia categories allowing the definition of a topic without having a priori knowledge of its domain. Furthermore, the model enables abstraction by accessing superordinate concepts not mentioned during conversation to define a dialog topic. By interfacing a conversational agent with the framework based on our model, we aim to improve human-agent conversations in that the artificial interlocutor is more topic-aware. Referring to this, the presented preliminary evaluation shows promising results.

The presented model implies many approaches to different challenges and some remain unexplained. Thus, we have several ideas for future work. First, we need to extend the evaluation in terms of other corpora more adequate for our purpose. Second, the graph-based representation of the Wikipedia taxonomy calls for the inclusion of graph-based similarity measures, e.g. introduced by Resnik [22]. Furthermore, we strive for utilizing the whole graph with all categories to enable the definition of a large variety of topics. The conceptualization and realization of an appropriate task specifying a probability threshold regarding the accuracy of a topic is also part of the future work. Eventually, we will employ the model to the architecture of our conversational agent.

ACKNOWLEDGMENT

This work is kindly supported by the Deutsche Forschungsgemeinschaft (DFG) in the context of the KnowCIT research project in the Center of Excellence Cognitive Interaction Technology (CITEC). We thank Kevin Fischer for assisting in the research and evaluation underlying this publication.

REFERENCES

- [1] R. Mitkov, Ed., *The Oxford Handbook of Computational Linguistics*. Oxford University Press, 2003.
- [2] J. Svennevig, *Getting Acquainted in Conversation*. John Benjamins Publishing, 1999.
- [3] 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*, 2011, (To appear).
- [4] W. Bublitz, "Topical coherence in spoken discourse," *Studia Anglica Posnaniensia*, vol. 22, pp. 31–51, 1989.
- [5] D. Goutsos, *Modeling discourse topic: sequential relations and strategies in expository text*. Ablex Publishing Corporation, 1997.
- [6] R. C. Schank, "Rules and topics in conversation," *Cognitive Science*, vol. 1, no. 4, pp. 421–441, 1977.
- [7] H. H. Clark, *Using language*, 7th ed. Cambridge Univ. Press, 2007.
- [8] J. Allan, *Topic Detection and Tracking: Event-based Information Organization*, J. Allan, Ed. Kluwer Academic Publishers, 2002.
- [9] J. Makkonen, H. Ahonen-Myka, and M. Salmenkivi, "Simple semantics in topic detection and tracking," *Information Retrieval*, vol. 7, no. 3–4, pp. 347–368, 2004.
- [10] X. Zhang and T. Wang, "Topic tracking with dynamic topic model and topic-based weighting method," *Journal of Software*, vol. 5, no. 5, pp. 482–489, 2010.
- [11] E. Gabrilovich and S. Markovitch, "Computing semantic relatedness using wikipedia-based explicit semantic analysis," in *Proceedings of the International Joint Conference on Artificial Intelligence*, 2007, pp. 6–12.
- [12] 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.
- [13] N. Eagle, P. Singh, and A. S. Pentlan, "Common sense conversations: Understanding casual conversation using a common sense database," in *Proceedings of the Artificial Intelligence, Information Access, and Mobile Computing Workshop (IJCAI 2003)*, 2003.
- [14] T. Bergstrom and K. Karahalios, "Conversation clusters: Grouping conversation topics through human-computer dialog," in *Proceedings of the Int. Conf. on Human Factors in Computing Systems (CHI09)*, 2009.
- [15] 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 Int'l Workshop on Robot and Human Interactive Communication (ROMAN)*, 2006, pp. 26–32.
- [16] K. Lagus and J. Kuusisto, "Topic identification in natural language dialogues using neural network," in *Proceedings of the 3rd SIGdial Workshop on Discourse and Dialogue*, 2002, pp. 95–102.
- [17] U. Waltinger and A. Mehler, "Social semantics and its evaluation by means of semantic relatedness and open topic models," in *Proceedings of the 2009 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, 2009, pp. 42–49.
- [18] M. Alemzadeh and F. Karray, "An efficient method for tagging a query with category labels using wikipedia towards enhancing search engine results," in *Proceedings of the IEEE/WIC/ACM Int. Conference on Web Intelligence and Intelligent Agent Technology*, 2010, pp. 192–195.
- [19] 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 Conf. on Empirical Methods in Natural Language Processing and Very Large Corpora*, 2000, pp. 63–70.
- [20] M. McCandless, E. Hatcher, and O. Gospodnetić, *Lucene in Action*, 2nd ed. Manning, 2010.
- [21] A. Downing, "Talking topically," *CIRCLE of Linguistics Applied to Communication (CLAC)*, vol. 3, pp. 31–50, 2000.
- [22] P. Resnik, "Using information content to evaluate semantic similarity in a taxonomy," in *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, 1995.