Computational Linguistics and Intellectual Technologies:
Proceedings of the International Conference “Dialogue 2020”

Moscow, June 17–20, 2020

DIALOGUE MANAGEMENT USING EXTENDED DISCOURSE TREES

Ilvovsky D. A. (dilvovsky@hse.ru)
Kazan Federal University, Kazan, Russia; National Research University Higher School of Economics Moscow, Russia

Galitsky B. A. (bgalitsky@hotmail.com)
Oracle Inc, USA

In this paper we learn how to manage a dialogue relying on discourse of its utterances. We consider two complementary approaches of dialogue management based on the discourse text analysis to extend the abilities of the interactive information retrieval-based chat bot.

Keywords: discourse tree, dialogue management, rhetoric structure

DOI: 10.28995/2075-7182-2020-19-361-371

УПРАВЛЕНИЕ ДИАЛОГОМ С ПОМОЩЬЮ РАСШИРЕННЫХ ДИСКУРСИВНЫХ ДЕРЕВЬЕВ

Ильвовский Д. А. (dilvovsky@hse.ru)
Казанский Федеральный Университет, Казань, Россия;
Национальный Исследовательский Университет
Высшая Школа Экономики, Москва, Россия

Галицкий Б. А. (bgalitsky@hotmail.com)
Оракл, США

Ключевые слова: дискурсивное дерево, риторические структуры, управление диалогом

1. Introduction

In this paper we extend the abilities of the interactive chat bot initially developed by [Galitsky and Ilvovsky, 2017] and later improved in [Galitsky, 2019]; [Galitsky and Ilvovsky 2019a]; [Galitsky and Ilvovsky 2019b]. We consider two complementary approaches to the dialogue management both using discourse analysis based on RST [Mann and Thompson, 1988] and. both utilizing Discourse Trees (DT) for the texts.

The first approach is inspired by an idea of a guided search. One source of it is a search methodology designed to show a user an array of different visual possibilities where a searching user may proceed. This is done instead of just navigating to an end point or a terminal answer. We believe that knowledge exploration should be driven by navigating an *extended* discourse tree (EDT) built for the whole corpus of relevant content. It is a combination of discourse trees of individual paragraphs first across paragraphs in a document and then across documents [Galitsky, 2019].

The second approach tries to force the user to request exhaustive information and explanation about the particular topic. We try to achieve this by utilizing discourse tree of the initial piece of a text. At each step of the conversation chat bot analyses remaining topics and try to make user more focused on the initial topic "turning" him back to the undiscussed parts of the text and forcing to request for more details.

As one can see these two approaches are complementary and can be easily combined in one interactive chat bot. In the rest of the paper we discuss these approaches in more details and provide preliminary evaluation for the second approach. More information about general chat bot architecture and evaluation can be found in [Galitsky and Ilvovsky, 2017]; [Galitsky, 2019].

2. Multi Document Navigation Using Extended Discourse Trees

To control the chat bot navigation in a general case, beyond clarification scenarios, one can use the notion of an **extended discourse tree** [Galitsky, 2019]. A conventional discourse tree expresses the author flow of thoughts at the level of paragraph or multiple paragraphs. Conventional discourse tree becomes fairly inaccurate when applied to larger text fragments, or documents. Hence we can extend the notion of a linguistic discourse tree towards an extended discourse tree, a representation for the set of inter-connected documents covering a topic. For a given paragraph, a DT is automatically built by discourse parsers [Joty et.al, 2014]. We then automatically combine DTs for the paragraphs of documents to the EDT, which is a basis of an interactive content exploration facilitated by the chat bot. We apply structured learning of extended DTs to differentiate between good, cognitively plausible scenarios and counter-intuitive, non-cohesive ones. To provide cohesive answers, we use a measure of rhetorical agreement between a question and an answer by tree kernel learning of their discourse trees [Galitsky and Ilvovsky, 2017].

On the web, information is usually represented in web pages and documents, with certain section structure. Answering questions, forming topics of candidate answers and attempting to provide an answer based on user selected topic are the operations which can be represented with the help of a structure that includes the DTs of texts involved. When a certain portion of text is suggested to a user as an answer,

this user might want to drill in something more specific, ascend to a more general level of knowledge or make a side move to a topic at the same level. These user intents of navigating from one portion of text to another can be represented as coordinate or subordinate discourse relations between these portions.

We merge the links between logical parts of paragraphs and the links between documents (Fig. 1). If at the current step the user is interested in drilling in, we navigate her through an *Elaboration* relation from nucleus to satellite within a paragraph or *Elaboration* hyperlink to a more specific document. Conversely, if a user decides that the suggested topic is not exactly what he is looking for and wants to return a higher-level view, the system navigates *Elaboration* relation *in the inverse order* from satellite to nucleus at either paragraph or intra-document level. The other navigation option is relying on *Contrast* or *Condition* relations exploring controversial topics (these rhetorical relations need to be recognized for inter-document case).

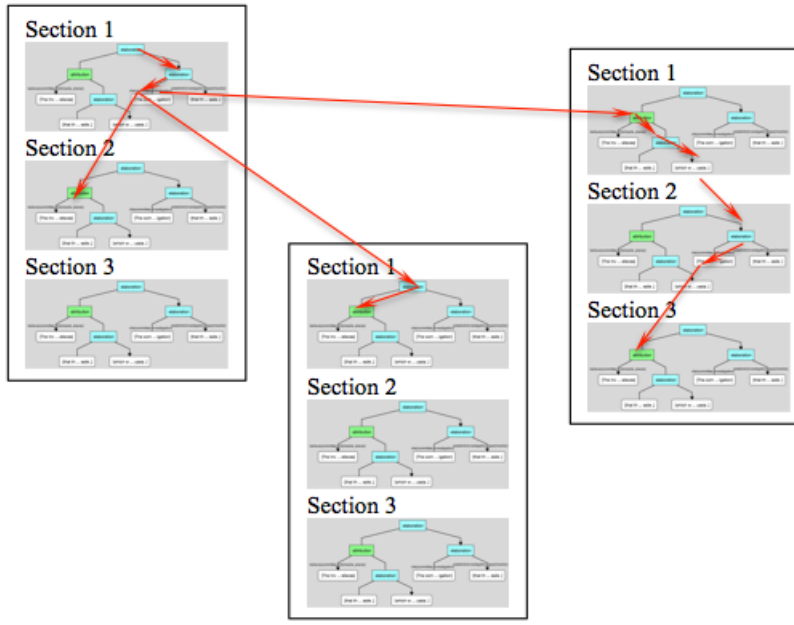

Fig. 1 Illustration for the idea of extended DT: intra-paragraph rhetorical relations are combined with inter-document links also labeled as rhetorical relations

Navigation starts with the route node of a section that matches the user query most closely. Then the chat bot attempts to build a set of possible topics, possible understanding of user intent. To do that, it extracts phrases from elementary discourse units that are satellites of the route node of the DT. If the user accepts a given topic, the navigation continues along the chosen edge; otherwise, when no topic covers the user interest, the chat bot backtracks the discourse tree and proceeds to the other section

(possibly of other documents) which matched the original user query second best. Inter-document and inter-section edges for relations such as $Elaboration$ play similar role in knowledge exploration navigation to the internal edges of a conventional DT.

3. Using Discourse Tree to Navigate a User through All Aspects of the Topic

3.1. How to Force a User to Request Exhaustive Information and Explanation

In many task-oriented chat bot domains, an objective is to fully inform a user about a particular important piece of information. It is also crucial to make user believe this piece of information, relying on explanation and argumentation in as much degree as possible. In some cases, it is important to make a user believe in a particular short text. This should be done by thoroughly navigating a user through possible disagreements and misunderstanding, to make sure the user is being explained and communicated an issue exhaustively.

1) If a text is given, navigating a discourse tree of this text T is one of the most efficient ways to communicate it. The chat bot starts with making an introduction I_T and then making the main statement M_T. Then the user would ask for more details E_T, disagree with the E_T or ask a question on a topic outside of the scope of this text O_T.
2) If the user asks for more details I_T, the EDU connected with $Elaboration$ with M_T is provided as a reply. We denote this EDU as $Elaboration(I_T)$. This is the easiest, most direct situation.
3) If the user disagrees, chat bot tries to find an EDU which is connected by Explanation or Cause with M_T or I_T. This EDU should be returned as a reply.
4) If the user asks a different question O_T then it should be answered as a factoid question but nevertheless the chat bot needs to take the user back to T so the reply should end with $Elaboration(I_T)$.
5) If the user doubts about the validity of a claim in M_T, the chat bot needs to deliver $Attribution(M_T)$ as an answer.

The procedure above should iterate until no more EDU in T is left or the user terminates the conversation. If the chat bot persistence is too high in trying to take the user back to T, this user would terminate the conversation too soon. Otherwise, if the chat bot persistence is too low, the user would deviate from T too far so will red less content of T (EDU(T)). We want to optimize the chat bot to maintain the optimal persistence to maximize the number of delivered EDU(T) till the conversation is abandoned by the user.

The chat bot can use different modes interacting with the user:

1) Always try to find an EDU matching the user query. Always give the closest EDU as an answer, even if it is very dissimilar to the question.

2) Try to find an EDU matching the user query. If a good match is not found then go to a foreign content. Then proceed with DT navigation. Give a foreign answer and then DT navigation EDU.
3) Try to find an EDU matching the user query. If a good match is not found then encourage a user to ask a question or make a comment which would match a rhetorical relation in a navigation path.
4) Try to search foreign content as much as possible.

3.2. Example of a Navigation using "Forcing" Approach

Initial Text: *According to BBC, China has rejected calls for an independent international investigation into the origin of the coronavirus. A top diplomat in the UK, Chen Wen explained the BBC the demands were politically motivated and would divert China attention from fighting the pandemic. However, EU believes that information about how it initially spread could help countries tackle the disease. The virus is thought to have been caused by a poor hygiene emerged at a wildlife market in the city of Wuhan.*[1]

Discourse Tree for the text

elaboration
  elaboration
    elaboration
      attribution
        TEXT: According to BBC,
        TEXT: China has rejected calls for an independent international investigation into the origin of the coronavirus.
      contrast
        joint
          explanation
            TEXT: A top diplomat in the UK , Chen Wen explained the BBC
            TEXT: the demands were politically motivated and would
            TEXT: divert China attention from fighting the pandemic .
          attribution
            TEXT: However, EU believes
            means
              TEXT: that information about how it initially spread could help
              TEXT: countries tackle the disease .
    cause
      TEXT: The virus is thought to have been caused by a poor
      TEXT: hygiene emerged at a wildlife market in the city of Wuhan.
  TEXT: Meanwhile, an EU report accuses China of spreading disinformation about the crisis

[1] https://www.bbc.com/news/world-asia-china-52420536

Sample dialogue with the technical comments

Bot: "China has rejected calls for an independent international investigation into the origin of the coronavirus"
User: Why would they do it? Don't believe that.
Bot: [doubt is detected] [(*Attribution* ("China has rejected…") = According to BBC] *This is* "According to BBC"
User: Are politicians involved in this?
Bot: [Searching for 'politi…' & 'involv…' tokens in EDUs] "the demands were politically motivated"
User: Why is that?
Bot: [Detected explainability request. Looking for *Explanation*("the demands were politically…".] "A top diplomat in the UK, Chen Wen explained the BBC
User: Who is Chen Wen?
Bot: [question is about an entity from the answer but foreign to the thread. So search on the web and then continue navigate DT] Chen Wen is … [Found information on the web]. *Notice also that* [*Elaboration, Joint*, or other default rhetorical relation from the current node] "and would divert China attention from fighting the pandemic."
User: And how can countries tackle coronavirus?
Bot: [Searching for 'countr…' & 'tackle…' & 'coronavirus' tokens in EDUs & rhetorical relations related to 'how' such as *Explanation, Cause, Means* etc]. Selected fragment:
EU believes
means
TEXT: that information about how it initially spread could help
TEXT: countries tackle the disease.
"EU believes that information about how it initially spread could help countries tackle the disease"
User: Why is it EU business?
Bot: [question is about an entity from the answer but foreign to the thread. So search on the web and then continue navigate DT]
[Found node with EDU: The virus is thought to have been caused … Formed question for this EDU and the relation]
Do you know what causes "the virus"
User: Yes I know, it is caused by …
Bot: [User confirmed. It does not matter what the user knows. Proceed to the next EDU & relation]
"It is caused by a poor hygiene emerged at a wildlife market in the city of Wuhan"

In *italic* we show the phrases added by the chat bot. In [square brackets] we show the results of chat bot analysis of utterances and a reply selection rule/result. EDUs from original text are shown in "double quotes".

What we observe is that a dialogue is fairly plausible, although no data-driven method has been applied. It does not matter where the user deviates from the target text as long as the chat bot always takes her back to the EDU and rhetorical relation which is either relevant to what the user asked or claimed, or just follows the DT navigation flow from more important, closer to the root, to less important. If the user is asked a question outside of the scope of the target text, we provide an answer from the foreign source and then also switch topic and come back to the business of the target answer, proceeding with the DT navigation.

Finally, we show just the **readable sample dialogue**:

Bot: China has rejected calls for an independent international investigation into the origin of the coronavirus
User: Why would they do it? Don't believe that.
Bot: This is According to BBC
User: Are politicians involved in this?
Bot: The demands were politically motivated
User: Why is that?
Bot: A top diplomat in the UK, Chen Wen explained the BBC
User: Who is Chen Wen?
Bot:
Chen Wen is …
and would divert China attention from fighting the pandemic
User: And how can countries tackle coronavirus?
Bot: EU believes that information about how it initially spread could help countries tackle the disease
User: Why is it EU business?
Bot: Do you know what causes the virus?
User: Yes I know, it is caused by …
Bot: It is caused by a poor hygiene emerged at a wildlife market in the city of Wuhan

4. Preliminary Evaluation

We tried to estimate how reasonable is to "force" a user to request more knowledge using our approach. We evaluated the overall meaningfulness of the utterances produced by the bot and average "level of user satisfaction".

The topics of the dialogues included product features, technical instructions, and latest news. Each discussion was based on the source text as it was shown in the example in Section 3.2. Five users (author's colleagues) interacted with the chat bot on these topics, and stopped when they think that dialogue is not useful for them anymore. After each conversation users assessed the meaningfulness of each bot utterance (just "yes"/"no") and then rated the whole "usefulness" of a dialogue using grades from 1 to 3 where 3 mean "dialogue was useful for me".

Table 1. The overall meaningfulness of the dialogues with the bot

Domain & source	News	Instructions	Product features
Number of dialogues	40	40	45
Average number of utterances	9	6	10
Percentage of meaningful bot utterances	0.73	0.69	0.78
Average user satisfaction	2.2	1.9	2.3

5. Related Work

Discourse disentanglement (such as classification of links between portions of texts or documents) and dialogue/speech/communicative act tagging have been extensively studied [Wang et al., 2011]. Discourse disentanglement is the task of splitting a conversation [Elsner and Charniak, 2008] or documents [Wolf and Gibson, 2005] into a sequence of distinct portions of text (sub-discourses). The disentangled discourse is modeled via a tree structure [Grosz and Sidner 1986]; [Seo et al., 2009], an acyclic graph structure [Rose et al., 1995]; [Elsner and Charniak, 2008], or a cyclic chain graph structure [Wolf and Gibson, 2005]. Speech acts are used to describe the function or role of an utterance in a discourse, similarly to our CDT representation, and have been employed for the analysis of communication means including conversational speech instant messaging, security analysis of documents [Galitsky and Makowski, 2017], online forums [Kim et al., 2010], [Galitsky et al., 2017] and chats [Galitsky and Ilvovsky, 2017a]. Automated answer scoring benefits from semantic and discourse analyses as well [Wanas et al., 2008]. For a more complete review of models for discourse disentanglement and speech act tagging, we refer the reader to [Kim et al., 2010].

[Wang et al. 2011] presented the task of parsing user forum threads to determine the labeled dependencies between posts. Three methods, including a dependency parsing approach, are proposed to jointly classify the links (relationships) between posts and the dialogue act (type) of each link. The authors predicted not only the links between posts, but also showed the type of each link, in the form of the discourse structure of the thread.

6. Conclusions and Future Work

We presented a discourse-based dialogue management system for a chat bot with iterative content exploration that leads a user through a personalized knowledge acquisition session. The chat bot is focused on automated customer support or product recommendation agent that assists a user in learning product features, product usability, suitability, troubleshooting and other related tasks.

Although there has been a substantial advancement in document-level RST parsing, including the rich linguistic features-based of parsing models [Joty et al., 2014], document level discourse analysis has not found a broad range of applications such as search. The most valuable information from DT includes global discourse features and long range structural dependencies between DT constituents.

A number of studies including [Surdeanu et al., 2015] showed that discourse information is beneficial for search. Our paper is explicitly showing how discourse trees help to navigate search. Discourse trees and their extensions is a very promising subject of study for logical AI. Logical AI studies subjects such as logic forms and logic programs which are very limited in quantity in the real world. But logical AI tries to make sense of them: discourse trees are fairly interpretable structures. Discourse trees can be obtained in large quantity on one hand and they are adequate Logical AI subject on the other hand. That is why discourse trees and their extension is such an important subject of study for search engineering and chat bots.

In question answering the current version of chat bot relies only to the data extracted from text documents. Now we are working on complementing these data by the data from Linked Open Data cloud, including DBpedia [Lehmann et al., 2015]. As an interface between natural language user query and LOD datasets we would rely on the resource from the Linguistic Linked Open Data cloud [Cimiano et al., 2020], such as LLOD representation of WordNet [McCrae et al., 2014], BabelNet [Ehrmann et al., 2014], FrameNet [Rospocher et al., 2019] and RuThes [Kirillovich et al., 2017]; [Galieva et al., 2017]. We expect that exploitation of LOD cloud can improve user's satisfaction against the baseline obtained in this work.

7. Acknowledgements

The new results on discourse-based navigation proposed in this paper were received by Dmitry Ilvovsky during his work in KFU and was funded by Russian Science Foundation according to the research project no. 19-71-10056. These results are based on the research in discourse models that were by made by Dmitry Ilvovsky and Boris Galitsky in NRU HSE during last years.

References

1. *Cimiano, P., Chiarcos, C., McCrae, J. P., and Gracia, J.* (2020). Linguistic Linked Open Data Cloud. In Cimiano, P., et al. Linguistic Linked Data: Representation, Generation and Applications, pages 29–41. Springer.
2. *Ehrmann, M., Cecconi, F., Vannella, D., McCrae, J., Cimiano, P., and Navigli, R.* (2014). Representing Multilingual Data as Linked Data: the Case of BabelNet 2.0. In Calzolari, N., et al., Eds., Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC 2014), pages 401–408. ELRA.
3. *Elsner, M. and E. Charniak.* (2008) You talking to me? a corpus and algorithm for conversation disentanglement. In Proceedings of the 46th Annual Meeting of the ACL: HLT (ACL 2008), pages 834–842, Columbus, USA.
4. *Galieva, A., Kirillovich, A., Khakimov, B., Loukachevitch, N., Nevzorova, O., and Suleymanov, D.* (2017). Toward Domain-Specific Russian-Tatar Thesaurus Construction. In Proceedings of the International Conference IMS-2017, pages. 120–124. ACM.

5. *Galitsky B. and G. Makowski* (2017) Document Classifier for a Data Loss Prevention System based on Learning Rhetoric Relations. CICLing 2017, April 17–23, Budapest, Hungary.
6. *Galitsky B. and McKenna E. W.* (2017) Sentiment Extraction from Consumer Reviews for Providing Product Recommendations. US Patent 9646078B2.
7. *Galitsky B., Ilvovsky D.* (2017) Chat bot with a Discourse Structure-Driven Dialogue Management, EACL Demo Program.
8. *Galitsky B., Ilvovsky D.* (2019b) On a Chatbot Conducting Virtual Dialogues. Proceedings of the 28th ACM International Conference on Information and Knowledge Management.
9. *Galitsky B., Ilvovsky D., Goncharova, E.* (2019a) On a Chatbot Conducting Dialogue-in-Dialogue. Proceedings of the 20th Annual SIGdial Meeting on Discourse and Dialogue.
10. *Galitsky B.* (2019) Developing Enterprise Chatbots – Springer International Publishing, 2019. Chapter 11. Discourse Level Dialogue Management.
11. *Galitsky, B., M. P. González, C. I. Chesñevar* (2009) A novel approach for classifying customer complaints through graphs similarities in argumentative dialogue. Decision Support Systems, Volume 46, Issue 3, pp. 717–729 (2009).
12. *Grosz B. J. and C. L. Sidner* (1986) Attention, intention and the structure of discourse. Computational Linguistics, 12(3):175–204.
13. *Joty S. R. and A. Moschitti* (2014) Discriminative Reranking of Discourse Parses Using Tree Kernels. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP).
14. *Kim, S. N., Li Wang, and Timothy Baldwin.* (2010) Tagging and linking web forum posts. CoNLL-2010, pages 192–202, Uppsala, Sweden
15. *Kirillovich, A., Nevzorova, O., Gimadiev. E., and Loukachevitch, N.* (2017). RuThes Cloud: Towards a Multilevel Linguistic Linked Open Data Resource for Russian. In Różewski, P. and Lange, C., Eds. Proceedings of the 8th International Conference on Knowledge Engineering and Semantic Web (KESW 2017). Communications in Computer and Information Science, vol. 786, pages. 38–52. Springer, Cham.
16. *Lehmann, J., Isele, R., Jakob, M., Jentzsch, A., Kontokostas, D., Mendes, P. N., Hellmann, S., Morsey, M., van Kleef, P., Auer, S., and Bizer, C.* (2015). DBpedia: A Large-scale, Multilingual Knowledge Base Extracted from Wikipedia. Semantic Web Journal, 6(2):167–195.
17. *Mann, W. and Thompson, S., Rhetorical Structure Theory: Toward a functional theory of text organization, 1988*
18. *McCrae, J P., Fellbaum, C., and Cimiano, P.* (2014). Publishing and Linking WordNet using lemon and RDF. In Chiarcos, C., et al., Eds., Proceedings of the 3rd Workshop on Linked Data in Linguistics (LDL-2014), pages. 13–16. ELRA.
19. *Rose CP, Barbara Di Eugenio, Lori S. Levin, and Carol Van Ess-Dykema* (1995) Discourse processing of dialogues with multiple threads. ACL.
20. *Rospocher, M., Corcoglioniti, F., and Palmero Aprosio, A.* (2019). PreMOn: LODifing linguistic predicate models. Language Resources and Evaluation, 53:499–524.

21. *Sakai, T.* (2007). Alternatives to Bpref. Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. Amsterdam, The Netherlands, ACM: 71–78.
22. *Seo, J., W. B. Croft, and D. A. Smith* (2009) Online community search using thread structure. CIKM 2009, pages 1907–1910, Hong Kong, China.
23. *Surdeanu M., T. Hicks, and M. A. Valenzuela-Escarcega.* Two Practical Rhetorical Structure Theory Parsers. NAACL HLT, 2015.
24. *Wanas, N., M. El-Saban, H. Ashour, and W. Ammar* (2008) Automatic scoring of online discussion posts. In Proceeding of the 2nd ACM workshop on Information credibility on the web (WICOW '08), pages 19–26, Napa Valley, USA.
25. *Wang, L., Lui, M., Kim, S. N., Nivre, J., Baldwin, T.*: Predicting thread discourse structure over technical web forums. EMNLP. pp. 13–25 (2011)
26. *Wolf, F. and E. Gibson* (2005) Representing discourse coherence: A corpus-based study. Computational Linguistics, 31(2):249–287.