Automatic planning of the dialogue between human and machine using discourse trees

Boris Galitsky **Dmitry Ilvovsky**

Oracle National Research University Higher School of Economics

In many task-oriented chatbot 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.

Rather than throwing the whole paragraph of text at a user, we split it into logical parts and feed the user text fragment by fragment, following her interests and intents. To systematically implement this navigation, we follow a discourse-level structure for how the author of this text organized his thoughts. This can be done by navigating a discourse tree (DT) of this text. DT is a tree that is a labeled tree in which the leaves of the tree correspond to contiguous units for clauses (elementary discourse units, EDUs). Adjacent EDUs, as well as higher-level (larger) discourse units, are organized in a hierarchy by rhetorical relation (e.g., *Reason, Temporal sequence)* provided by Rhetorical structure theory (RST, Mann and Thompson, 1988). An anti-symmetric relation involves a pair of EDUs: nuclei, which are core parts of the relation, and satellites, which are the supportive parts of the rhetorical relation. A satellite can be delivered by the chatbot to a user as an utterance only if its nucleus has already been received and acknowledged in one way or another.

We outline the chatbot algorithm of the DT traversal, covering a multitude of user intents at each iteration:

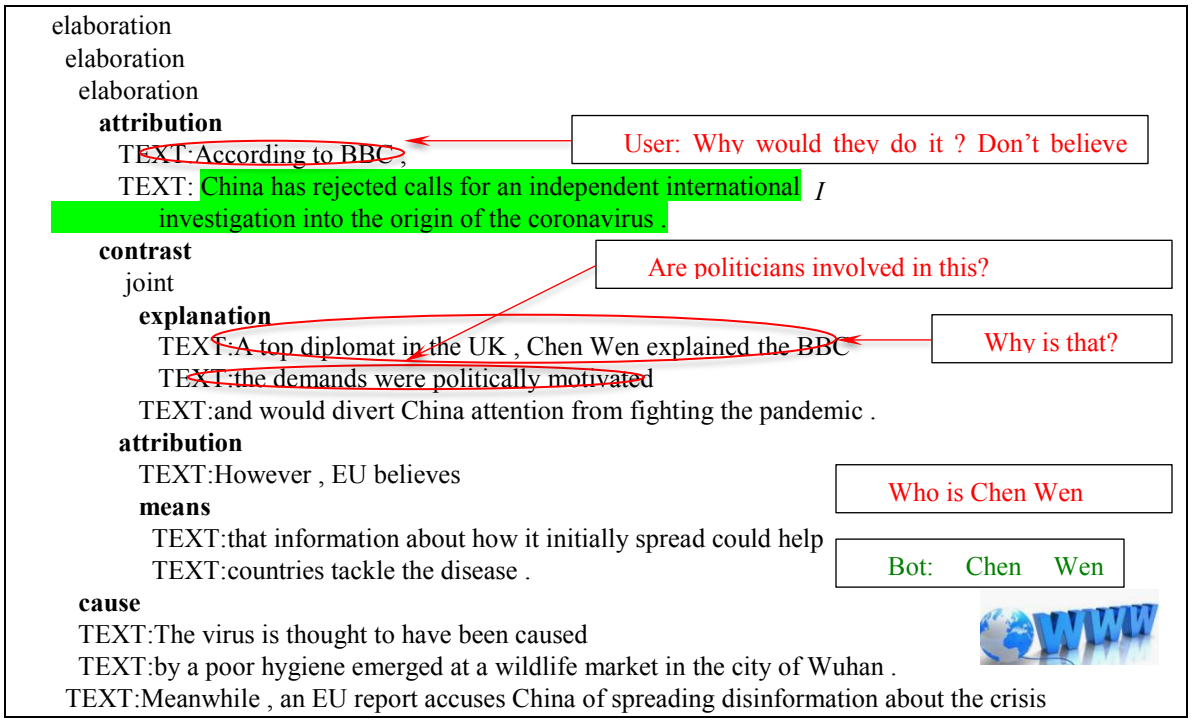
1. If a text is given, navigating a discourse tree of this text T is one of the most efficient ways to communicate it. The chatbot starts with making an introduction and then making the main statement M_T. Then the user would ask for more details E_T, disagree E_T or ask a question on a topic outside 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, chatbot 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 chatbot 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 chatbot needs to deliver *Attribution*(M_T) as an answer.

The procedure above should iterate till no more EDU in T is left or the user terminates the conversation. If the chatbot persistence is too high in trying to take the user back to T, this user would terminate the conversation too soon. Otherwise, if the chatbot 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 chatbot to maintain the optimal persistence to maximize the number of delivered EDU(T) till the conversation is abandoned by the user.

Let us take a text and show how a DT navigation leads a dialogue wrapped around this 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.

A discourse tree for this text and a fragment of a sample navigation path is shown below.

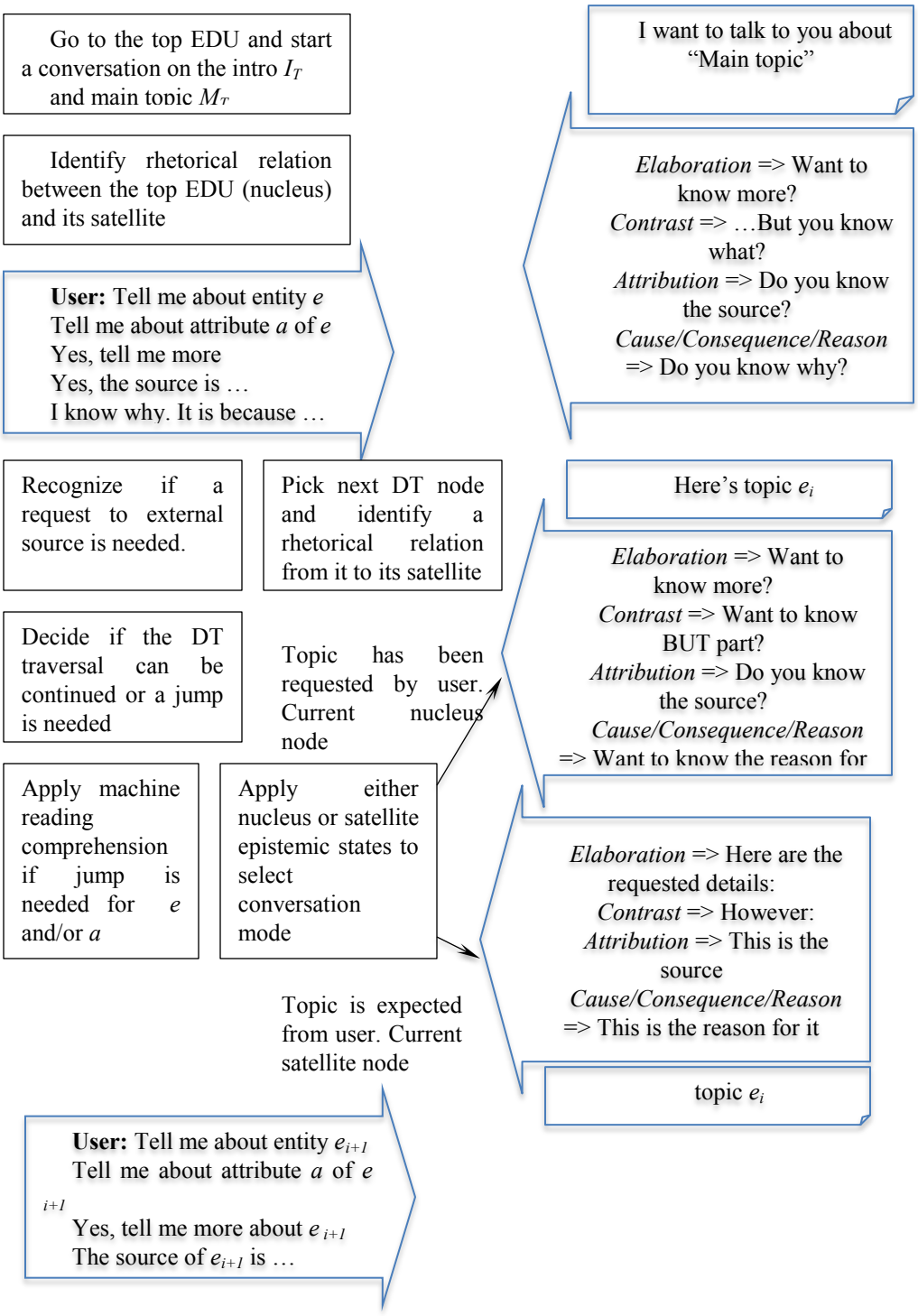


What we observe is that a dialogue is fairly plausible, although no data-driven method has been applied. It does not matter where the user deviates from the target text as long as the chatbot always takes her back to the EDU and rhetorical relation which is either relevant to what the user asked or claimed, 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.

The dialogue flow based on navigation of a DT is shown below. A conversation with the focus on T starts with an Introduction of T followed by the main topic of T expressed by EDUs located closest to the root of DT. Chatbot utterance includes the information from the EDU of the current node plus an encouragement to the user to continue conversation, such as a question or a knowledge sharing request (on the top-right). Chatbot encouragement depends on the rhetorical relation for the current navigation node (now, the M_T node). The user replies (formulates a question) in a certain form, depending on the encouragement question of the chatbot (on the left).

User's question varies in terms of the focus entity or its attribute, and/or the epistemic state initiated by the chatbot. Once the user question is received by the chatbot, it is analyzed with respect to if an external knowledge source needs to be searched and/or if an Machine Reading Comprehension (MRC) component needs to be initiated to find a value for a factoid question and also identify an EDU this value occurs in. Then the decision needs to be made if the user changed the topic and a jump is required, or the chatbot can maintain the dialogue by continuing the DT navigation. Next navigation step depends on whether the current node is nucleus (and satellite is the next to be visited), or it is a satellite and its nucleus needs a visit. Epistemic state update is chosen accordingly.

For the nucleus, the user has already expressed his interest in a given topic. So information from its EDU is ready to be sent to the user. For the satellite, the user is encouraged to express his interest according to the rhetorical relation to the nucleus of this satellite. A topic is expected from this user. External search and/or MRC can be applied in this option.



References

Mann, W. and Thompson, S., Rhetorical Structure Theory: Toward a functional theory of text organization, 1988