Imperial College London

SUPPLEMENTARY DOCUMENT FOR MSCI PHYSICS REPORT

IMPERIAL COLLEGE LONDON

BLACKETT LABORATORY

Discussion Trees: Visualising Conversation

Author:
CID: 01348239

Project Code:
THEO-Evans2

Supervisor:
Dr. Tim Evans

Assessor:
Dr. Kim Christensen

Date: April 20, 2021

1 Introduction

This document is an extension to my MSci project report, **Natural Language Processing Methods for Conversation Analysis and Visualisation**. In my main report I discuss how we developed two state-of-the-art Natural Language Processing Methods for Conversation Analysis, namely the Bi-GRU-CRF model for Dialogue Act classification, and the Graph of Embedded Extracted Keywords (GEEK) algorithm for topic extraction. I also define **Discussion Trees**, visualisations of conversation transcripts built to highlight the structure and topical evolution of human-human discussions. In this report I will describe the process through which Discussion Trees were developed, as well as a discussion on the most exciting future application of Discussion Trees: Business Meeting (BM) Discussion Trees.

Before settling on the final Discussion Tree structure, a few preliminary visualisations were developed, namely the Trajectory Through Topic Space, the Stepping Line, and the Shifting Topic Step. Each of these will now be discussed.

2 The Trajectory Through Topic Space

I was interested in how topics are localised or change for individual speakers in a conversation, and as such, the first method of conversation visualisation investigated in this project was the **Trajectory Through Topic Space**. Topic space is defined as the layout of the ConceptNet Numberbatch word vectors (1) for each of the topical keywords. As speakers move from one topic to another, a quiver line is drawn between the nodes representing the topic keywords. An example of a Trajectory Through Topic Space is given in Fig. 1.

Prior to the use of ConceptNet Numberbatch word embeddings, two other pre-trained word embeddings were trialled: Word2Vec's `GoogleNews-vectors-negative300` and FastText's `cc.en.300.bin` (2). The distribution of keywords was visually most accurate when Numberbatch embeddings were used and hence these were chosen for the final Trajectory Through Topic Space graphics. I also experimented with plotting a separate trajectory for the utterances spoken by different participants in the conversation, an example of which is given in Fig. 2.

2.0.1 Limitations

When all utterances from a conversation are considered, the appearance of the Trajectory Through Topic Space method of conversation visualisation becomes too intricate; a viewer is not able to decipher the temporal change in topic and the trajectory hides information on the topics being discussed. I therefore decided to stop experimenting with this style of visualisation, and take my approach back to basics.

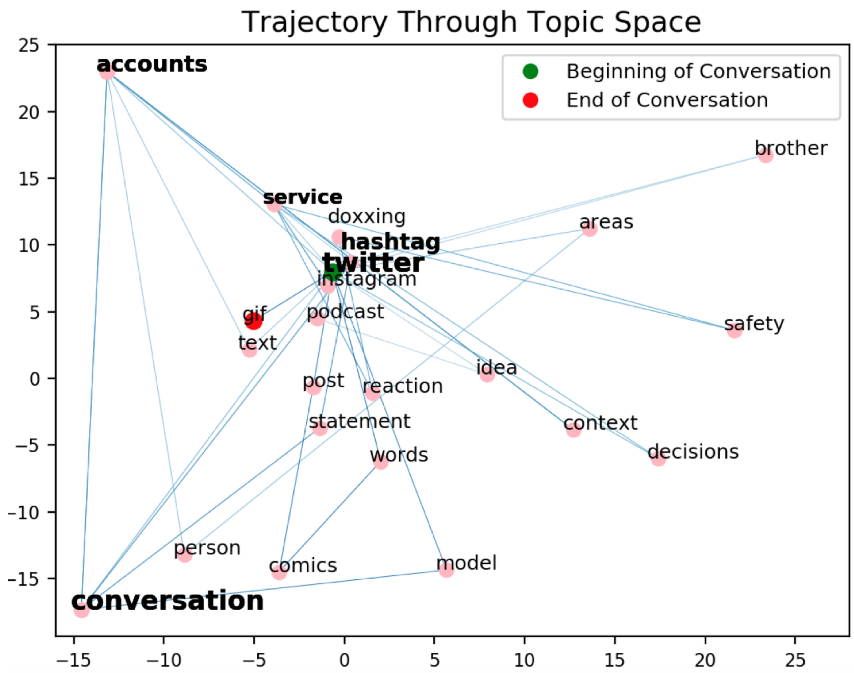

Figure 1: A Trajectory Through Topic Space of the first 400 utterances in the Joe Rogan interview of Jack Dorsey. The size of plotted words reflects the number of times they were mentioned during the podcast as a whole. Note that although axis values have been left in this visualisation they have very little meaning and hence are removed from here forward.

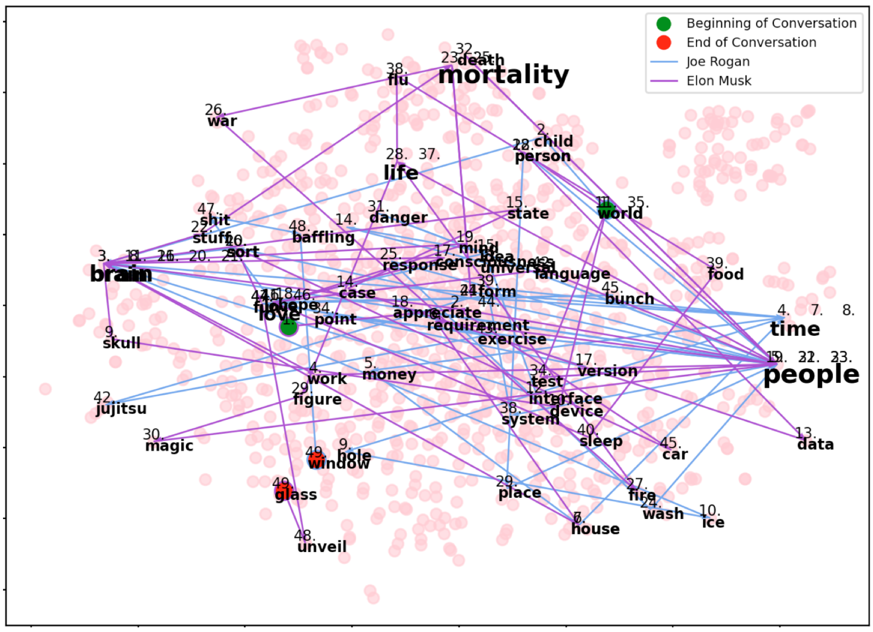

Figure 2: The speaker-wise Trajectory Through Topic Space for the interview of Elon Musk on the Joe Rogan Podcast show. The trajectory of Joe Rogan's utterances is depicted in blue; the trajectory of Elon Musk's utterances in purple. The size of plotted keywords reflects their usage in the transcript as a whole.

3 The Stepping Line

Having evaluated the previous visualisation method and decided the appearance was much too complex to allow for easy information extraction by a viewer, a new approach to visualisation was required. I decided to begin by investigating how topics change and who introduced them and for this I used a simple 'Stepping' line graphic.

3.1 Do Questions Introduce Topics?

I began with the hypothesis that it is the use of *questioning* Dialogue Acts[1] which leads the topical evolution of a conversation. To test this, I created the **Stepping Question Line**: a simple visualisation wherein a (black) node is plotted one step along the x-axis for every speaker turn, with the trajectory shifting one step in the +y direction if a question was asked by a speaker during their speaker turn. The node is then annotated with the topic currently under discussion (extracted using our GEEK algorithm). An example of this is given in Fig. 3, where only the first 200 utterances of the conversation between Joe Rogan and Elon Musk are considered. By colouring the line to reflect the current speaker we can see that almost all of the questions asked in the first 200 lines of the transcript are asked by Joe Rogan, as is expected given he is the interviewer, but that these steps don't seem to align with changes in topic.

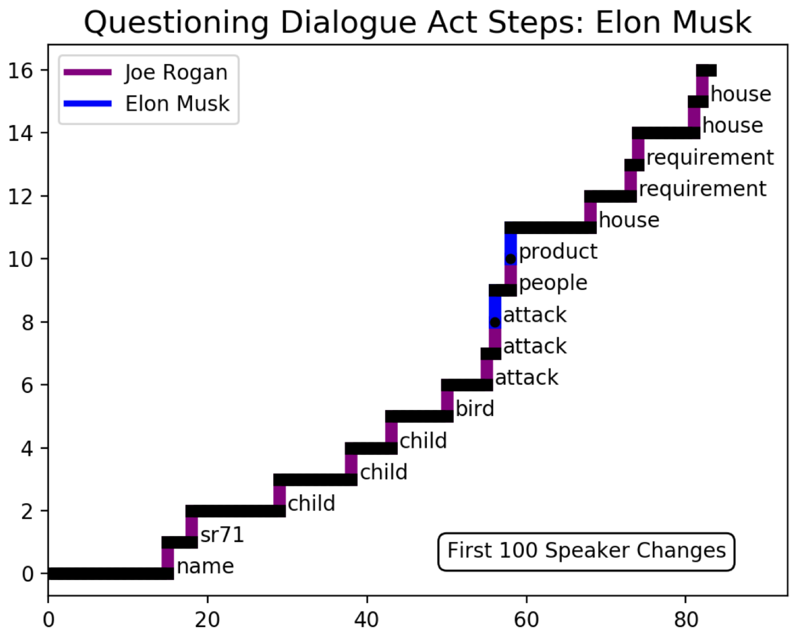

Figure 3: Questioning-Dialogue Act Step visualisation of the transcript of the Joe Rogan podcast interview of Elon Musk. Horizontal steps are taken for each speaker turn; vertical steps are taken when a question is asked during a speaker turn. Axes have very little meaning and are removed in all further visualisations.

[1]Questioning Dialogue Acts include: "Wh-Question", "Yes-No-Question", "Declarative Yes-No-Question", and "Declarative Wh-Question".

3.2 Who Introduces Topics?

To investigate whether the speaker who asks the most questions is also the speaker who introduces the most topics, I created the **Stepping Topic Line** visualisation. The only difference between this and the previous visualisation is that here I am instead shifting up the line when a *change in topic* is detected within the given utterance. I used this visualisation to test the hypothesis that it is the interviewer who introduces all the key topics in a conversation by again colouring the line according to which speaker first mentions each new topic.

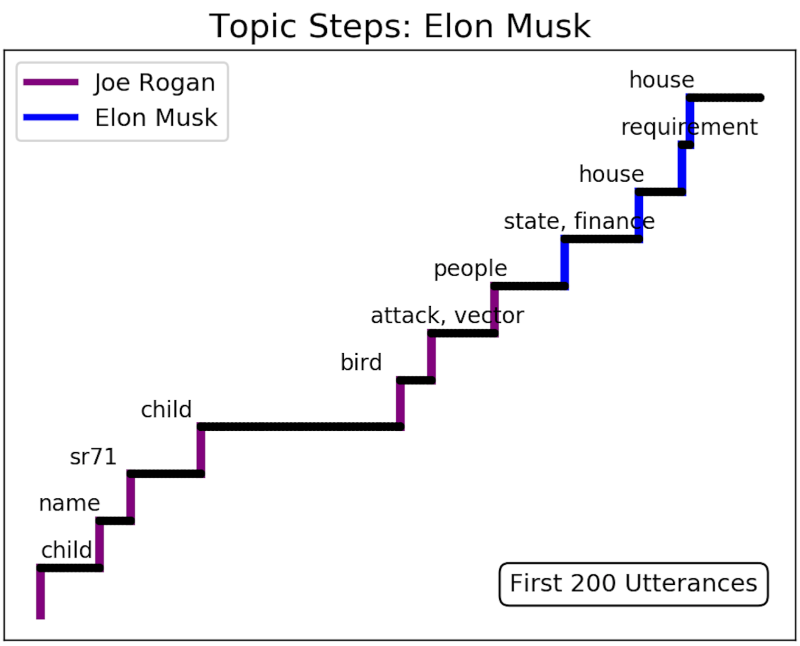

Figure 4: Stepping Topic Line visualisation of the Joe Rogan podcast interview of Elon Musk transcript. Horizontal steps are taken for each utterance; vertical steps are taken at a change of conversation topic. Only the first 200 utterances are considered here for visual clarity. Axis labels are removed as they have very little meaning.

From Fig. 4 we see that Joe Rogan introduces most of the topics in the first 200 utterances of the conversation – which again makes sense as he is the interviewer and hence will be introducing topics to get Elon Musk settled in – however within these first few utterances there is again no clear pattern regarding who introduced the most topics. I therefore decided to turn to developing a visualisation which highlights a conversation's *structure*, rather than fine details.

4 The Shifting Topic Step

The first change I made to the Stepping Topic Line visualisation shown in Fig. 4 was to change the axis along which nodes are stacked when their corresponding utterances continued to speak about the same topic, namely from the x-axis to the y-axis. This meant that there was now a unique x-axis position for every topic mentioned during the conversation. The next modification made was the introduction of horizontal shifts back to the relevant x-positions for every time the conversation looped back to a previously-mentioned topic. This modification is shown in Fig. 5a where the first 300 utterances of the same Joe Rogan Interview of Elon Musk are plotted. The visualisation is then cleaned up in Fig. 5b by removing the speaker colour and only labelling topics which were mentioned multiple times during the conversation. This **Shifting Topic Step** visualisation illustrates the natural loops which occur in conversations as participants link back to previously-discussed topics.

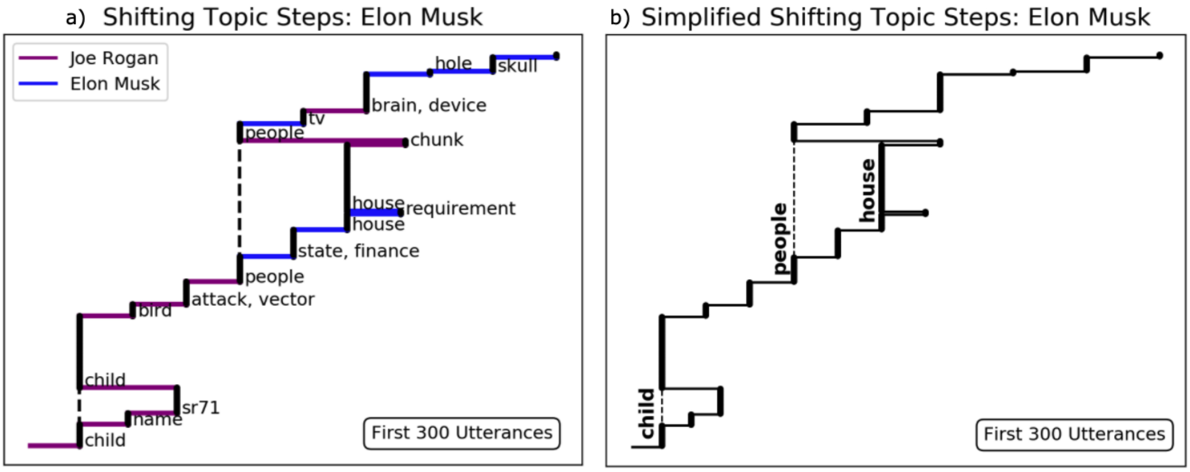

Figure 5: Illustrations to aid the explanation of how the Shifting Topic Step visualisation was developed. Analysis done on the first 300 utterances in the Joe Rogan interview of Elon Musk.

4.1 Developing the Discussion Tree...

Three key modifications were then made to this Shifting Topic Step visualisation to develop it into the first Discussion Tree structure. First, instead of *shifting* the line every time a previously-mentioned topic was brought up again in conversation, I instead *started a new branch* from the relevant x-position. This update can be seen in Fig. 6 wherein I have also annotated the 'branch number' on the orange nodes which mark the position of the last utterance of every branch. To further clean up the appearance and move towards a more tree-like structure, I alter the graphic such that each time a new branch begins, the x increments are sent a different way along the x axis: this modification is shown in Fig. 7. Finally, I shift the branches down such that they begin from the top of the stack of utterances which most recently spoke about that topic. Finally, we have a basic **Discussion Tree** structure, shown in Fig. 8.

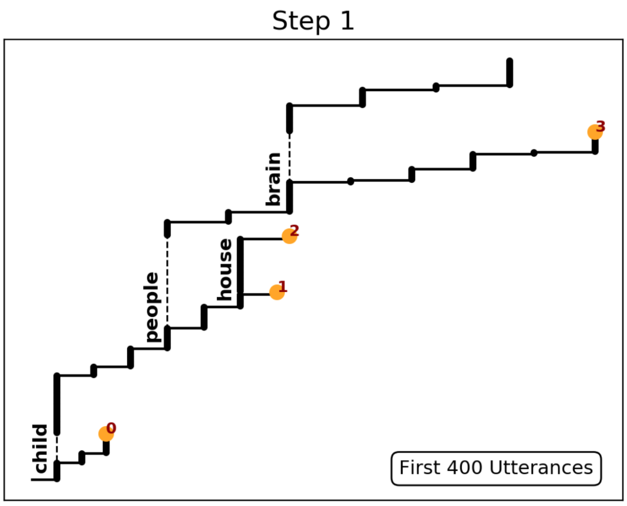

Figure 6: To develop the visualisation in Fig. 5b, a new branch is started at the relevant x-position for a topic which is looped back to in the conversation. The topics which are looped back to are annotated here, along with the branch number in red (plotted at the position of the last utterance spoken about the relevant topic).

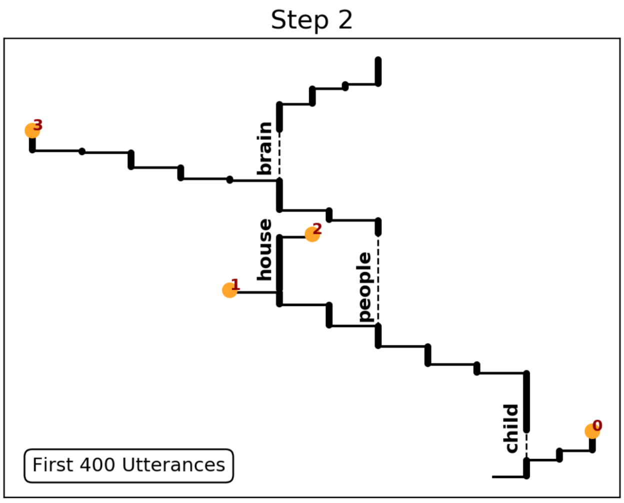

Figure 7: To clean up the appearance of the visualisation, the branches are sent in opposite directions along the x-axis. Branch numbers are again annotated in red at the end of each branch; the highest branch has no label as it was cut off by us only considering the first 400 utterances of the transcript. As a reminder, all visualisations in this section are built from the transcript of the Joe Rogan interview of Elon Musk.

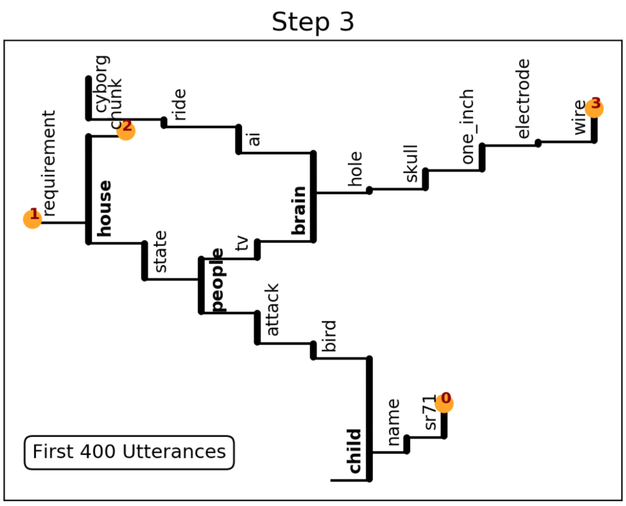

Figure 8: The final step in the development of a basic Discussion Tree structure was for the branches to be shifted down such that they begin from one step above the node representing last utterance which spoke about the relevant topic. Here, all stack labels are plotted, with those in bold representing topics looped back-to during the first 400 utterances of this transcript.

5 Discussion Trees

5.1 Tree Structure

For the sake of completeness, I now re-define the key structural features of Discussion Trees:

- **Branches** are lines of fully-connected nodes which represent consecutive utterances in a conversation.
- **Stacks** are vertically-aligned sections of a branch, built from consecutive utterances on the same topic. There are often multiple stacks within a branch of a conversation.
- **Leaves** are the coloured nodes, plotted to highlight the last utterance on each branch. In this project they are presented in just one colour for the sake of visual clarity however in future they could be used to provide further information about the given branch.

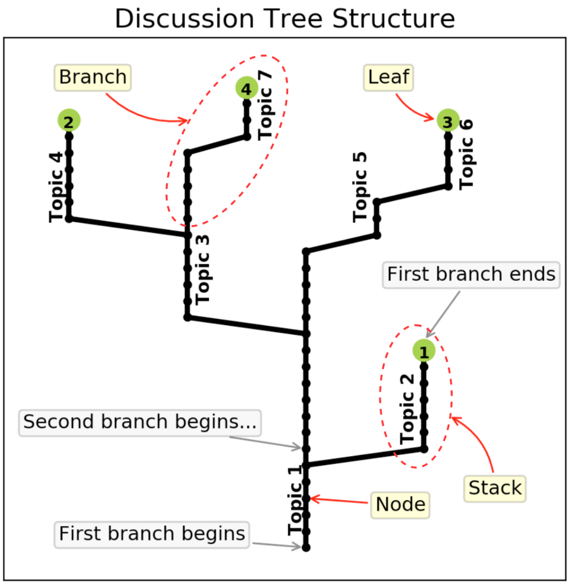

Figure 9: The basic structure of a Discussion Tree, with key tree features annotated in red, and explanatory annotations given in grey. Individual utterances from the transcript are plotted as black nodes; vertically stacked nodes indicate consecutive utterances on the same topic labelled in black, and horizontal shifts indicate changes in conversation topic. Branches end when the conversation jumps back to a previously-mentioned topic. Leaves are plotted on the last node of each branch, with the relevant branch number annotated. The full definitions of stacks, branches, and leaves are given in 5.1.

The decision was made to remove all axis labels due to the numbers having very little meaning. The value in these graphics is revealed when tree structures are placed side-by-side: it is their *relative* structure and size of features that sheds light on the nature of the given conversation, and less so their exact dimensions. For example, the relative length of each stack indicates the amount of time spent discussing that topic and, in general, a Discussion Tree with many long branches indicates that the speakers covered many new topics in a 'linear' fashion whereas a Discussion Tree with many short branches indicates that the conversation kept looping back to the same topics.

5.2 Final Comments

We present Discussion Trees: visual fingerprints of conversation transcripts which utilise the output of our GEEK algorithm to build a macroscopic view of conversations whose structure offers a new way to generate insights on the topical evolution and overall nature of human-human interactions. Fig. 10 presents the Discussion Trees for nine episodes from the 'Heavy Topics' Spotify Podcast show, illustrating how such graphics can be used as visual fingerprints for each episode.

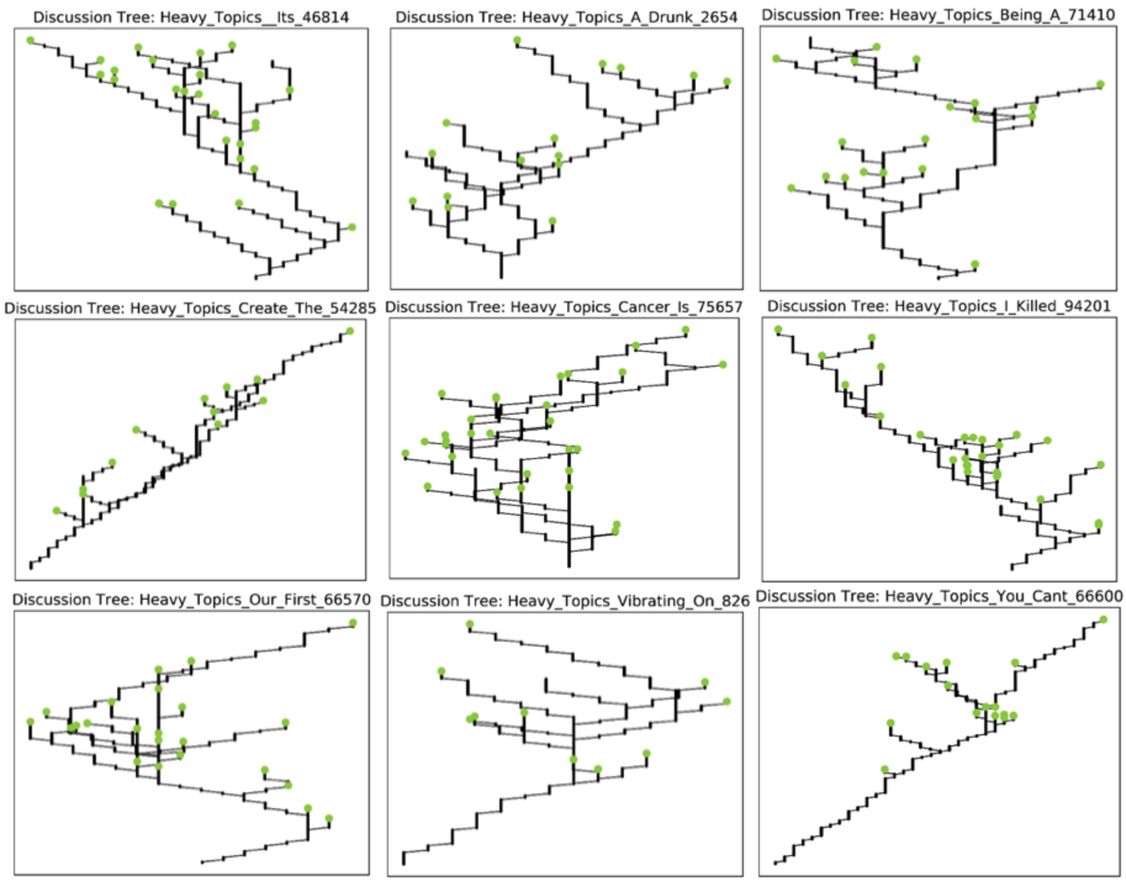

Figure 10: The Discussion Trees for nine episodes from the 'Heavy Topics' podcast, available on Spotify. The trees are presented here to illustrate how they can be used as visual fingerprints for episodes of a given podcast.

6 Business Meeting (BM) Discussion Trees

6.1 The Vision

We believe there is a clear opportunity for the use of Discussion Trees in formal discussion environments such as political debates and business meetings. In particular, we propose **Business Meeting (BM) Discussion Trees**: Discussion Trees designed for use in corporate group discussions.

BM Discussion Trees will be designed such that they provide meeting participants and organisers with a sleek visual take-away from each meeting; a graphic which makes the key topics discussed accessible, highlighting which points need to be built out and which have yet to be explored, all at just a glance. An example of how the Discussion Tree visualisation could be incorporated into an automatic transcription app is given in Fig. 11.

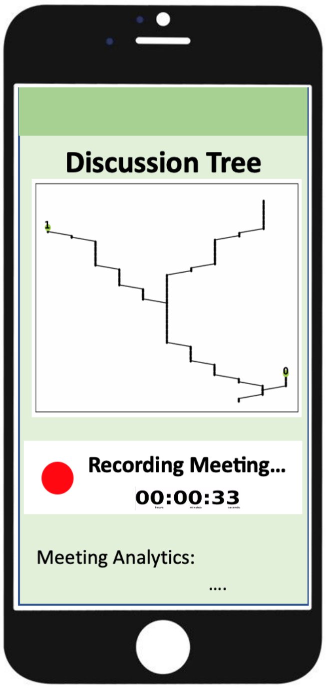

Figure 11: A very basic illustration to highlight how Discussion Trees could be incorporated into an automatic transcription smartphone application. The Discussion Tree would be built over time as the conversation progresses. Once the meeting ends every participant in the meeting would be sent a copy of the graphic; it would offer them a quick summary of the topics discussed during the meeting, as well as a concise list of conclusions reached and to whom any tasks were assigned. The basic Discussion Tree appearance would be updated to better suit this commercial application according to the modifications discussed in section 6.2.

6.2 Updated Framework

In order to update the appearance of the basic Discussion Tree structure such that it provides more specialised information for Business Meeting participants, I propose a few key modifications to the original Discussion Tree framework. The first is to define two types of stacks for BM Discussion Trees:

1. **To-Be-Discussed (TBD) stacks**: stacks corresponding to topics which are present in a to-be-discussed topic list that the organiser sets prior to the meeting taking place. An ideal meeting would only cover these TBD topics and would reach an

actionable conclusion at the end of each one. TBD-stacks would be coloured in black on the Discussion Tree.

2. **Unnecessary Chat (UC) stacks**: stacks whose topic are *not* part of the To-Be-Discussed topic list and are therefore classed as Unnecessary Chat. Note that topics automatically classified as UC topics could be classed as TBD topics after a meeting has ended by the meeting organiser. TBD-stacks would be coloured in grey on the Discussion Tree.

A slight modification to the GEEK algorithm would also be implemented, namely that before proceeding with the usual decision-making process for topic segment and label extraction, GEEK would compare the keywords in each clump of utterances with all topics on the TBD list - if the cosine difference is large enough it would then automatically assign the TBD topic as the cluster label. This would likely results in multiple clusters of utterances –which GEEK initially thought represented different topics but which actually represented sub-topics of the larger umbrella TBD-topic – being joined together and hence fewer stacks will exist in the discussion tree overall. This would benefit the visualisation as it would clean up the appearance of the tree and leave more room for labels.

To visualise the conclusions the group comes to when discussing each TBD topic, we define two possible types of leaves, again differentiated by colour:

1. **Green Leaves**: for discussions which reach a definite conclusion, i.e. "This idea will not go ahead", or "Yes, we will begin this project on Monday".

2. **Amber Leaves**: for when groups are unable to reach a conclusion, i.e. "I will have to check with Stacey from the communications department before we proceed".

As the Discussion Trees would have fewer stacks and branches than the examples given in Fig. 10, we could add these leaves to the end of any TBD stack, rather than only to the ends of branches. Participants of the meeting, or meeting organisers, could therefore quickly observe the *Actionable Points*[2] from the meeting, and to whom they were assigned. To detect the above conclusion types we would hand-label a number of meetings, or access an appropriate dataset if there is one, and train a neural network to perform the classification task.

We hypothesis that the tree structures will distinguish productive meetings from unproductive meeting. For example, a well-structured meeting should cover each TBD topic one at a time and in a sufficient amount of detail to reach a conclusion. Therefore a DT of a perfectly productive meeting would move only between TBD-stacks with only one main branch; whereas an unproductive meeting which contains side-chatter would have many more stacks which represent not-TBD topics, and the group may have to return to certain topics if they were not discussed in enough

[2]Tasks raised in a meeting that require further work – often immediately assigned to a certain member of the group.

detail the first time and this would create a great number of branches. Therefore, we hypothesise that the structure of Discussion Trees for productive meetings will be long, slim, and mainly coloured in black (see section 6.2), whereas unproductive and inefficient meetings will have a bushy structure with the majority of stacks coloured in grey. We could therefore define a scalar metric for the *efficiency* of a meeting based simply on the fraction of stack coloured in grey or black (colouring introduced in section 6.2).

7 Further Extensions

Within this MSci project, two state-of-the-art methods of Natural Language Processing for Conversation Analysis were developed. Namely, these were the Bi-GRU-CRF model for Dialogue Act Classification, and the Graph of Embedded Extracted Keywords (GEEK) algorithm for topic extraction. A couple of extensions to these methods are now discussed.

A property of conversations that could next be explored would be the **uncertainty** about a topic. Research into annotating uncertainty in spoken language has been undertaken (3; 4); the Polson and Curtis (2010) (5) paper shows explicitly how uncertainty evolves during a group conversation in a formal meeting environment. Discussion Trees could highlight the uncertainty about a topic using line thickness such that it is clear which topics need to be revisited and which conclusions reached need to be scrutinised; this representation could be used by participants to plan follow up work.

Transcripts miss out significant amounts of **non-verbal information** contained in human-human conversations such as the body language of speakers, their eye movements, physical actions, and tone (6), all which are expressions of human attitude. Therefore, an important extension of this research would be the inclusion of information extracted from multi-modal data to the transcript analysis seen here. To achieve this, the ICSI meetings dataset (7) could be used as these meetings were carried out in an intelligent meetings room and hence multi-modal data is available alongside every meeting transcript.

References

[1] R. Speer, J. Chin, and C. Havasi, "ConceptNet 5.5: An open multilingual graph of general knowledge," 2017, pp. 4444–4451. [Online]. Available: http://aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14972 pages 2

[2] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, "Enriching word vectors with subword information," **CoRR**, vol. abs/1607.04606, 2016. [Online]. Available: http://arxiv.org/abs/1607.04606 pages 2

[3] H. Pon-Barry, S. Shieber, and N. Longenbaugh, "Eliciting and annotating uncertainty in spoken language," in **Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)**. Reykjavik, Iceland: European Language Resources Association (ELRA), May 2014, pp. 1978–1983. [Online]. Available: http://www.lrec-conf.org/proceedings/lrec2014/pdf/1167_Paper.pdf pages 12

[4] J. Dral, D. Heylen, and R. op den Akker, "Detecting uncertainty in spoken dialogues: an exploratory research for the automatic detection of speaker uncertainty by using prosodic markers," in **Affective Computing and Sentiment Analysis**. Springer, 2011, pp. 67–77. pages 12

[5] D. Polson and A. Curtis, "Dynamics of uncertainty in geological interpretation," **Journal of the Geological Society**, vol. 167, no. 1, pp. 5–10, 2010. pages 12

[6] T. Angert and B. Schneider, "Augmenting Transcripts with Natural Language Processing and Multimodal Data," in **Proceedings of the 12th International Conference on Educational Data Mining**. International Educational Data Mining Society, 2019. pages 12

[7] E. Shriberg, R. Dhillon, S. Bhagat, J. Ang, and H. Carvey, "The ICSI meeting recorder dialog act (MRDA) corpus," in **Proceedings of the 5th SIGdial Workshop on Discourse and Dialogue at HLT-NAACL 2004**. Cambridge, Massachusetts, USA: Association for Computational Linguistics, Apr. 30 - May 1 2004, pp. 97–100. [Online]. Available: https://www.aclweb.org/anthology/W04-2319 pages 12