Analysis of Topic Propagation in Therapy Sessions Using Partially Labeled Latent Dirichlet Allocation

Ilyas Chaoua[1], Sergio Consoli[2(✉)], Aki Härmä[2], Rim Helaoui[2], and Diego Reforgiato Recupero[1]

[1] Mathematics and Computer Science Department, University of Cagliari, Via Ospedale 72, 09124 Cagliari, Italy
ilyaschaoua@gmail.com, diego.reforgiato@unica.it

[2] Philips Research, High Tech Campus 34, 5656 AE Eindhoven, The Netherlands
{sergio.consoli,aki.harma,rim.helaoui}@philips.com

Abstract. The full comprehension of how topics change within psychotherapeutic conversation is key for assessment and therapeutic strategies to adopt by the counselor to the patients. That might enable artificial intelligence (AI) approaches to recommend the most suitable strategy for a new patient. Basically, understanding the topics dynamics of previous cases allows choosing the best therapy to perform for new patients depending on their current conversations.

In this paper we leverage Partially Labeled Dirichlet Allocation with the goal to detect and track topics in real-life psychotherapeutic conversations. On the one hand, the detection of topics allows us identifying the semantic themes of the current therapeutic conversation and predicting topics ad-hoc for each talk-turn between the patient and the counselor. On the other hand, the tracking of topics is key to understand and explore the dynamics of the conversation giving insights and tips on logic and strategy to adopt.

We point out that the entire conversation is structured and modeled according to a sequence of ongoing topics that might propagate through each talk-turn. We present a new method that combines topic modeling and transitions matrices that gives important information to counselors for their therapeutic strategies.

Keywords: Conversational AI · Psychotherapeutic conversations · Topics detection and modeling · Partially Labeled Dirichlet Allocation · Transitions matrices

1 Introduction

Analysis and research of therapeutic conversations is a growing domain of research: technological advances and innovations in areas such as Natural Lan-

Authors are listed in alphabetic order since their contributions have been equally distributed.

F. Koch et al. (Eds.): AIH 2018, LNAI 11326, pp. 59–75, 2019.
https://doi.org/10.1007/978-3-030-12738-1_5

guage Processing (NLP), Semantic Web, Big Data, Artificial Intelligence and healthcare have triggered such a growth [13].

Humans interactions have been targeted and machine learning approaches have been employed to detect unusual patterns. An example, authors in [6] aimed to infer predictive models to structure task-oriented dialogs.

Cognitive Behavior Therapy (CBT) entails therapeutic conversational methods which consist of therapies that aim at treating mental health problems, emotional challenges, sleeping difficulties, relationship problems, drug and alcohol abuse, anxiety and depression. Such therapies tackle and try to change the way of thinking and behaving of patients. Figure 1 shows a diagram of the CBT rationale[1].

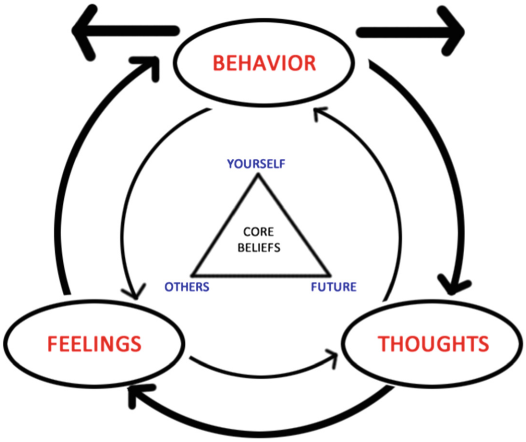

Fig. 1. The diagram depicts how emotions, thoughts and behaviors are related to each other. The inner triangle represents CBT's tenet that all humans' core beliefs can be summed up in the three mentioned categories.

More in detail, these therapeutic methods work by changing people's attitudes and their behavior by focusing on the thoughts, images, beliefs and attitudes that are held (a person's cognitive processes) and how these processes relate to the way a person behaves, as a way of dealing with emotional problems. The treatment is relatively short, taking five to ten months for most emotional problems. Patients usually attend one session per week where each session lasts less than 1 h. It is usually a face-to-face interaction between the counselor and the patient where the former needs to understand patient's feelings, e.g. confident, anxious, or depressed, as well as the causes of his feelings.

The conversation consists of a series of spoken sentences. Each is characterized by a certain topic: this creates a thematic structure to the whole therapeutic

[1] Image taken from Wikipedia https://en.wikipedia.org/wiki/Cognitive_behavioral_therapy.

conversation. The counselor employs techniques and strategies coming from clinical practice: he/she reacts to the patient and drives the conversation towards certain themes in order to tackle and solve the psychological problems of the patient.

It follows that the analysis and modeling of these human-to-human dialogues may be useful for the development of AI-based dialogue systems able to recommend the most appropriate therapeutic strategy to adopt by the practitioner for a new patient [6]. In such a context, a subset of the NLP research is related to the problem of topic detection and tracking (TDT), which has been widely focused and studied and combined with AI methods in the literature [15]. One goal of the TDT is the identification of the new topics in a conversation and their reappearance. Authors of [12] provide extensive background about that.

In this paper we focus on the conversation between counselor and patient and model the propagation of the topics identified during a given therapeutic conversation by using Partially Labeled Latent Dirichlet Allocation (PLDA) [21].

Traditional Latent Dirichlet Allocation (LDA) is one of the most popular topic model in the literature. LDA is based on a bag-of-words approach and is a generative statistical model that allows sets of observations to be explained by unobserved groups that explain why some parts of the data are similar. Our choice reflected on PLDA because the data we have used in our study are partially labeled and PLDA tends to achieve higher precision than traditional LDA on them.

The dataset we have used consists of 1729 real-life transcribed psychotherapeutic conversations, each made of different talk-turns. Further details about the data will be given in Sect. 4. Our approach works as it follows:

- First, we identify the most common topics used within the dataset.
- Then, the PLDA model takes as input the given conversations and detects significant words for each topic.
- The trained PLDA model is thus able to determine the potential topic addressed in each talk-turn.
- Within each conversation, the talk-turns flow is then transformed into a sequence of potential topics.
- Finally, for each topic, the semi-supervised PLDA topic model is evaluated by computing its coherence over the most significant words.

Our ultimate goal is to detect the key patterns within therapeutic conversations and to identify the topic switches according to the adopted dialogue strategy and topics propagation dynamics. In our method we are able to distinguish the topic changes driven by the counselor and the ones prompted by the patient. Two topic transition matrices are constructed accordingly to evaluate the two different topic changes. These matrices provide a numerical summary of the conversation and can be exploited to obtain tips for the overall understanding of the topics propagation dynamics.

This paper is further structured as it follows. Related works and literature background on automatic topic detection methods and therapeutic dialogue analysis are discussed in Sect. 2. LDA-based topic modelling algorithms are presented in Sect. 3, where we describe how they work, specifying their main characteristics and giving details about them. Section 4 includes details of the dataset we have employed for the experiments and the preprocessing we have performed on that. Our approach on TDT is presented in Sect. 5. Section 6 includes details related to the evaluation of our approach. Finally Sect. 7 ends the paper with conclusions and directions where we are headed.

2 Related Work

There is earlier work in the area of computational analysis of therapeutic sessions using topic modeling techniques, see, e.g., [5,17]. However, there seems to be few earlier works on dynamics of topic propagation in those conversations.

Digitalizing spoken interactions and recommending specific treatments are two current trends within the therapeutic conversations research. They use effective NLP technologies with the goal of extracting knowledge in text form from consultation transcripts. For example, work in [3] discussed an idea to combine communication theory used in healthcare and a visualization text analytic technique called *Discursis*, with the goal of analyzing the conversational behavior in consultations. More specifically, Discursis[2] is a computer-based tool for analysing human communication that can assist practitioners in understanding the structure, information content, and inter-speaker relationships that are present within input data. Discursis processes conversation transcript data to determine the conceptual content of each conversation turn. It offers visualizations and reports on the above information such as a concept map of the communication content, communication channels, concept recurrence matrix, score cards for each conversation in terms of, e.g., leader, follower, innovator, promoter.

The classification of conversations is not an easy task in medical consultations as it includes intense performance requirements, it is time-consuming and it suffers from non-standardized annotating systems. Authors of [14] presented an automated annotating system which leverages a Labeled LDA model [20] to assess the relationships between a certain conversation and its annotations. Annotations are related to the subjects symptoms present within the therapeutic conversations. The system is therefore able to identify the relevant annotations in separate talk-turns.

LDA also been used in a different way by authors in [16]. In particular, they have analysed a LDA topic model [8] as an automatic annotator tool for the topics and therapy prediction of the conversation. One assumption made by the authors was related to the fact that the automated detection of topics can be used to predict factors such as patient satisfaction and ratings of the therapy quality rather than predict the symptoms. The employment of the Labeled LDA and the LDA indicates that the identification and tracking of topics can provide important information to clinicians. They can use such information to better assist the patients and improve their treatments.

[2] http://www.discursis.com/.

Other authors analysed human communications in [2] where they developed a discourse visualization system which converted transcribed conversations to time series, a text analysis model and a set of quantitative metrics to detect and assess significant features. Their system was able to identify the topics adopted in a certain discussion by the participants and generate reports for each conversation.

These metrics can be seen as an extension of recurrence quantification analysis into the symbolic domain. The proposed technique may be used to monitor the state of a communication system and inform about interaction dynamics, including the level of topic consistency between participants; the timing of state changes for the participants as a result of changes in topic focus; and, patterns of topic proposal, reflection, and repetition.

Other researchers proposed one more use of the LDA model. In particular they adopted a conceptual dynamic latent Dirichlet allocation (CDLDA) model for TDT in conversational text content [24]. The differences between traditional LDA and the CDLDA model is that whereas the former employs bag-of-words techniques to identify topics, the latter considers information such as speech acts, semantic concepts, and hypernym definitions in E-HowNet [11][3]. The proposed method extracts the dependencies between speech acts and topics, where hypernym information makes the topic structure more complete and extends the abundance of original words. Results performed by the researchers proposing this idea indicated that the approach outperforms the conventional Dynamic Topic Models [7], LDA, and support vector machine models, achieving very high performance for TDT.

Work performed in [1] includes OntoLDA for the task of topic labeling. OntoLDA adopts an ontology-based topic model and a graph-based topic labeling method. Basically, the topic labeling method is based on the ontological meaning of the concepts included in the discovered topics. This approach indicated each topic as a multinomial distribution of concepts, and each concept as a distribution of words. OntoLDA scaled better the topics coherence score than the classical LDA. This was achieved by combining ontological concepts with probabilistic topic models towards a combined framework applied to various types of text collections.

One more approach that improved the human-agent dialogs was presented by authors in [9]. Their approach leveraged the basis of contextual knowledge provided by Wikipedia category system. To build their approach they had to map the different utterances to Wikipedia articles and define their relevant Wikipedia categories as a list of topics. It followed that the detection method was able to recognize a topic without holding a priori knowledge of its subject category.

Authors in [16] questioned the use of LDA to cast more light on the role of topic modeling to provide a measure of content more general than word features with the goal to identify patient satisfaction and evaluations of therapy quality. The unsupervised model they introduced produces models similar to manual annotation, and it appears to be better at predicting evaluations of the therapeutic relationship and important features of communication style particularly

[3] http://ckip.iis.sinica.edu.tw/taxonomy.

that of the counselors. This may suggest that unsupervised models used in this way are able to discover and track topics to provide more insight to therapists, enabling them to better direct their conversations in time-limited consultations, and serve the identification of patients who can afterward be at risk of relapse or non-adherence to treatment.

3 Topic Modeling

Topic models are a family of probabilistic approaches that aim at discovering latent semantic structures in large documents. Based on the presumption that meanings are relational, they interpret topics or themes within a set of documents originally constructed from a probability distribution over words. As a result, a document is viewed as a combination of topics, while a topic is viewed as a blend of words.

One of the most widely used statistical language modeling for this end is Latent Dirichlet Allocation (LDA) introduced by Blei et al. [8]. LDA is a generative approach. It assumes that documents in a given corpus are generated by repeatedly picking up a topic, then a word from that topic according to the distribution of all observed words in the corpus given that topic. LDA aims at learning these distributions and inferring the (hidden) topics given the (observed) words of the documents [18]. Given the nature of our data which includes partial annotations, we employ the following two variants of LDA.

Labeled Latent Dirichlet Allocation (LLDA) [20] is a supervised version of LDA that constraints it by defining a one-to-one correspondence between topics and human-provided labels. This approach, illustrated in the probabilistic graphical model in Fig. 2, allows LLDA to learn word-label correspondences.

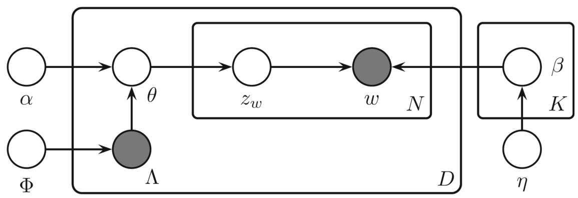

Fig. 2. Probabilistic graphical model of LLDA: unlike standard LDA, both the label set Λ as well as the topic prior α influence the topic mixture θ. β represents a vector of the parameters of the multinomial distribution whereas η are the parameters of the word prior. ω is the word, z is the per-word label assignments and ϕ is the label prior. Please check [20] for further details.

Partially Labeled Latent Dirichlet Allocation (PLDA) [21] is a semi-supervised version of LDA which extends it with constraints that align some learned topics with a human-provided label. The model exploits the unsupervised learning of topic models to explore the unseen themes with each label, as well as unlabeled themes in the large collection of data. As illustrated in Fig. 3, PLDA assumes that the document's words are drawn from a document-specific mixture of latent topics, where each topic is represented as a distribution over words, and each document can use only those topics that are in a topic class associated with one or more of the document's labels. This approach enables PLDA to detect extensive patterns in language usage correlated with each label.

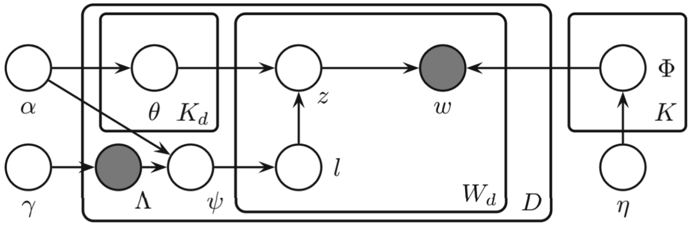

Fig. 3. Probabilistic graphical model for PLDA: each document's word ω and label Λ are observed, with the per-doc label distribution ψ, per-doc-label topic distributions θ, and per-topic word distributions Φ hidden variables. Because each document's label-set λ_d is observed, its sparse vector prior γ is unused; included for completeness. η are the parameters of the word prior whereas l is a label and z a topic. For further details please check the work in [21].

4 Experimental Dataset

In this section, we describe the dataset used in our experiments and the applied preprocessing steps. The used dataset consists of a collection of psychotherapeutic transcripts available for research. The transcribed and collected conversations adhere to the guidelines of the American Psychological Association (APA)[4]. An approval to use the collection was granted by an Internal Committee of Biomedical Experiments (ICBE) of Philips after a review of the agreements, the consent procedures, and data handling plan by legal and privacy experts. We remark that meta-data preprocessing has been executed for the two tasks we present in this paper (TDT), whereas text preprocessing has been run on topic detection only.

[4] http://www.apa.org.

4.1 Data Description

Counseling and psychotherapy transcripts contained in our data encompass a diverse set of patients, a large-scale array of topics, and different therapeutic strategies. Hundreds practicing counselors worldwide have transcribed and rendered the used conversations according to APA Ethics Guidelines for use and anonymity. The dataset consists of 1729 transcripts of 1:1 conversation with a total of 340,455 talk turns, 75,732 unique terms, and more than 9 million words. Each transcript has on average 200 talk-turns and eight words for talk-turn. They are also extended with meta-data consisting of the corresponding school of psychotherapy, counselors-patients information such as gender, age range, and sexual orientation as well as a table of topics discussed during the therapeutic conversation. Two different kind of information is contained in the table of topics, that is:

– *Subjects*, which are specified hierarchically into three consecutive levels. The top level is the most general subject, whereas the remaing two levels are more precise[5]
– *Symptoms*, which are overall 79 symptoms defined in the DSM-IV[6] manual, like, for example, *Depression, Anger, Fear*. Reason why we have adopted the DSM-IV and not the newer DSM-V is because the dataset we have employed has been structured according to DSM-IV.

4.2 Preprocessing of the Meta-data

Considering the high number of items in the table of topics, similar topics have been merged experimentally by means of the following steps:

1. Eliminate all the subjects and symptoms that occur in less than 3% of the dataset;
2. Group together all the subjects belonging to the same Wikipedia category[7] regardless their position in the given hierarchical structure.
3. Assign a label to the new subject according to the psychology topics table from APA. For example *Parent-child_relationship* and *Family* are mapped to a new subject from APA known as *Parenting*.
4. Reduce the number of symptoms by using the DSM-IV manual with the expert support of a counselor. In particular, we group symptoms with high-level correlation into a representative one. For example, *Sadness* and *Hopelessness* are merged into the symptom: *Depression*.

In this way the final set, illustrated in Fig. 4, has been reduced to 18 subjects and 16 symptoms only.

[5] For example, the word *Family* could correspond to a top level topic, while *Family violence* and *Child abuse* would be associated to the second and third levels respectively. Up to 575 subjects have been used in the three levels in total.

[6] https://dsm.psychiatryonline.org.

[7] https://en.wikipedia.org/wiki/Category:Main_topic_classifications.

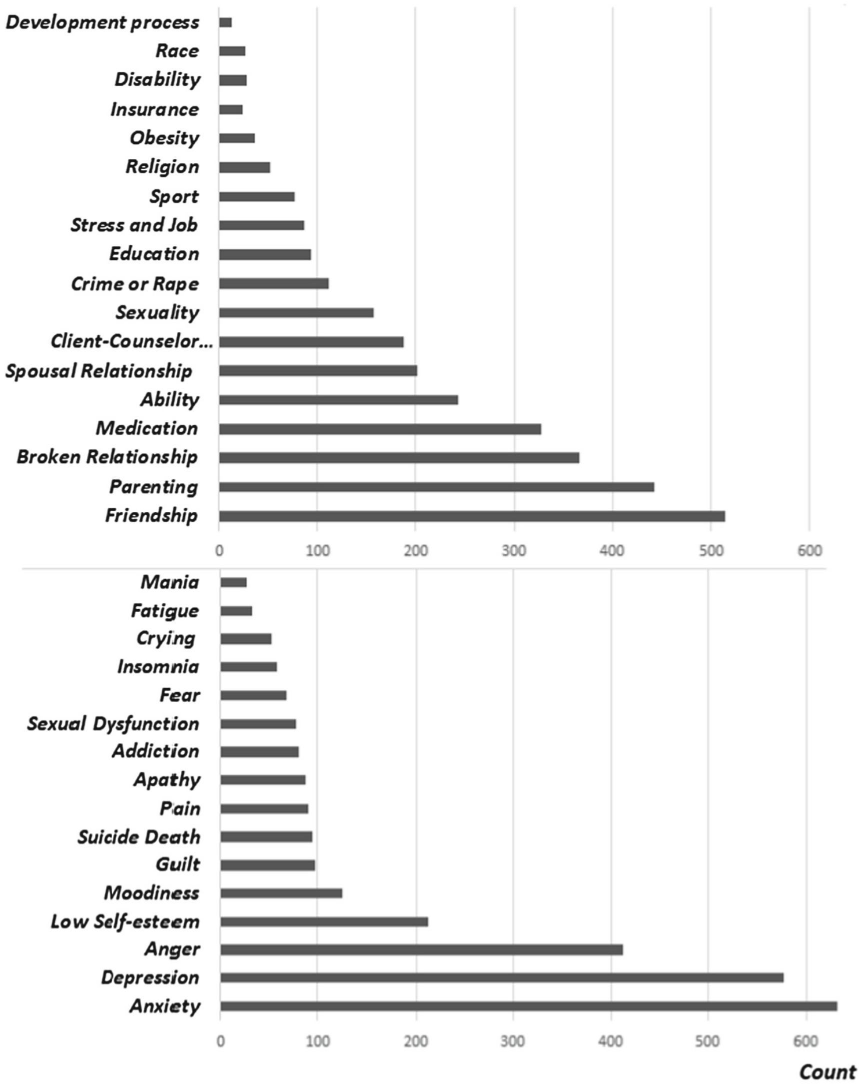

Fig. 4. The resulting 18 subjects (up) and 16 symptoms (bottom)

4.3 Preprocessing of the Conversation Text

A number of classic NLP pre-processing steps [23] have been applied to the dataset by using the NLTK platform[8]. The performed steps include:

1. tokenization, which transforms texts into a sequence of tokens;

[8] http://www.nltk.org/.

2. removal of all punctuations, stop words, numbers, words that frequently appeared in the text with minor content information (e.g., "mm-hmm"), and words that occurred in less than five documents whereas keeping nouns, verbs and adjectives only. This step was achieved by using unigram part-of-speech tagging [19], contained in the NLTK framework, in order to identify word types in each talk turn.
3. removal of the most common words (100 overall, in our case), and of the talk turns with one word only or with words shorter than three characters.

Note that the stemming and lemmatization steps have been omitted on purpose to avoid negative impact because the resulting changes may influence the evaluation of the topic model.

After the performed pre-processing steps, the resulting corpus consists of 2,849,457 tokens (14,274 unique ones) and a total of 268,478 talk turns.

5 The Proposed Approach for TDT

As described below, our proposed method consists of three phases; (1) topic modeling, (2) assignment of topic labels to talk turns, and, finally (3) tracking of the propagation of topics over the conversation.

5.1 Topic Modeling

The topic detection was performed using a PLDA implementation based on the Stanford Topic Modeling Toolbox[9](TMT). The model requires a set of parameters including the number of hidden topics to be discovered, the α and η hyperparameters (see Fig. 3), and a training corpus. We define each talk-turn as a document. This results in a total of 268,478 documents after the reprocessing step. Each document is associated with the corresponding topics from the table of topics of the corresponding transcript. Additionally to the 34 predefined topics inferred from the metadata, we experimentally set the number of hidden topics to 20. The hyperparameters α and η are set to 0.01. Based on those parameters, we train our model with 150 epochs using an approximate variant of the collapsed variational Bayes algorithm or the so called VB0 algorithm [4].

5.2 Topic Inference

The learned model delivers a weighted set of words per topic as illustrated in Table 1. The table depicts the top ten terms for each topic. The first column shows an example of a latent (i.e. discovered topic) whereas the second and third columns show two predefined topics together with their related words. As expected, terms associated to a particular topic tend to be semantically related particularly for subjects and symptoms. For example, the topic *Parenting* is defined by terms including family members, such as *mom, mother,*

[9] https://nlp.stanford.edu/software/tmt/tmt-0.4/.

dad, etc. whereas the topic *Addiction* includes terms close to alcohol and drugs (*drinking, smoke, etc.*). The same holds for latent topics, where *Topic-5*, for instance, includes similar terms related to *work*. Based on the learned weighted lists of terms, the model infers the most likely topics given a particular talk turn. Figure 5 depicts the *per-document topic distribution* of the five most likely topics in a selected conversation (i.e. *Stress and Job*; *Suicide and Death*; *Sexuality*; *Depression*; *Fear*).

Table 1. An illustrative example of the three kinds of topics and their most likely associated terms. *Topic-5* shows an example of the discovered topic, *Parenting* presents an example of a known subject, and *Addiction* presents an example of a known symptom.

Discovered topic: *Topic-5*		Known subject: *Parenting*		Known symptom: *Addiction*	
Associated words	Weight	Associated words	Weight	Associated words	Weight
Pay	1125386	*Mom*	1354.692	*Drinking*	120.9583
Month	957.7061	*Mother*	1190.696	*Drugs*	78.25677
Working	809.5385	*Dad*	1103.559	*Alcohol*	60.09412
End	799.9208	*Family*	996.8899	*Drug*	59.39509
Help	649.7758	*Brother*	786.8062	*Stoned*	47.30778
Giving	516.6524	*Parents*	745.5274	*Smoking*	44.19726
Months	502.375	*Father*	689.4923	*Marijuana*	41.16274
Year	463.4732	*Sister*	490.4897	*Girlfriend*	37.31469
Paid	421.3631	*Kids*	376.6678	*Smoke*	36.22964
Paying	416.7115	*Children*	313.5798	*Uptight*	35.90694

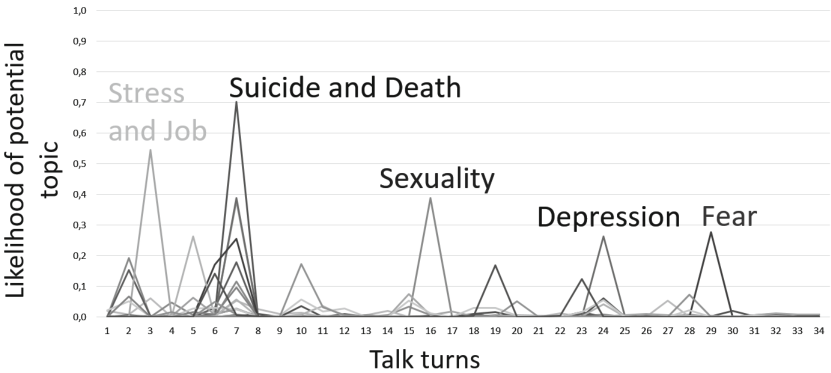

Fig. 5. Example of the *per-document topic distribution* in each talk-turn on a conversation.

Table 2 reports some examples of talk-turns of the patient and their associated representing topics, with the corresponding probabilities, produced by our PLDA-based method.

Table 2. Examples of talk-turns and their associated topics after PLDA within different conversations

Talk-turns	Associated topic
I came into it late, and it was a story about a father and daughter. And it was very much about feelings and...this was a man whose only family was his daughter and...had reappeared in her life and all that. And I remember thinking "Oh I bet Dad's not watching this at all." Or...is not enjoying it because I don't think ever of my family could feel ever be shared. Not with mom and me but even there I mean there were layers of...constraint. Etc.	Parenting, 40%
No. I don't know-maybe I do like it underneath it all. You know it keeps coming back to this climaxing - that I think I would enjoy intercourse if I could climax. And that seems to be you know, know-maybe know keeps coming climaxing think to enjoy intercourse climax seems to know	Sexual dysfunction, 92%
It's a kind of close friendship I guess of being able to just talk to them about anything or to not talk to them about anything. I mean just to sort of be able to be with them and have them understand how your feeling if you happen to be feeling any way at all or do enjoy things with you. Etc.	Friendship, 46%
Right. These feelings. That you know beginning to wonder if you know. You know I'm going to be this unhappy in marriage. To feel this lonely in the marriage. I don't want to be alone you know. In fact, I have all of the responsibility but none of the advantages. I want just to know have some of the advantages of being alone. And it feels pretty screwed up	Spousal relationship, 42%
Like sometimes when I'm thinking about sex or just getting away from everything including the person I'm talking to, and I don't feel like I can say that to a person right to his face	Sexuality, 82%
I never get really happy about anything very rarely, and at the same time, I never get really depressed about anything. I just don't let myself you know. And I was consciously sitting there trying to get - I mean after I started getting depressed I decided to relax and get just as depressed as I could get because Meg says that often helps	Depression, 46%
It's craving the marijuana. It's craving the alcohol. It's craving you know whatever it is	Addiction(s), 99%
All right sure. What effect does the medications we have you on now which is predominantly Lamictal and we have you on some Trazodone at night for sleep and I understand that s a catch 22 type of medication	Medication, 76%

5.3 Tracking Topics

In order to investigate how topics propagate throughout the given therapeutic conversations, we built two topics transition matrices (TTMs). For convenience, we added a new topic annotated as *Meaningless talk* which we associated to talk-turns that provide poor semantics contents or language, or non-verbal communication (e.g. "Yahh!!, Mm-hmm"). In general, the three main types of topics changes are depicted:

1. The counselor keeps talking about the same topic to the patient from the previous talk-turn, and vice versa;
2. The counselor moves to a new topic after the talk-turn of the patient;
3. The patient moves to a new topic after the talk-turn of the counselor.

To illustrate the topics dynamics in more details, we create a patient-to-counselor TTM CP_k that describes all the topic changes within the conversation k. In particular, $CP_k[i, j]$ is the number of times that topic i changes into topic j in a conversation k. We merge the CP_k matrices together by summing them up and obtain our final matrix CP. Similarly, we built a counselor-to-patient matrix PC by using the topics-change defined earlier by switching counselor and patient. The arithmetic difference between the two matrices is illustrated in Fig. 6.

The resulting values are mapped into colors to visually indicate the different levels of engagement between the counselor and the patient: black corresponds to values below -10, grey to values between -10 and $+10$, and greater white for values greater than 10. Thus, the diagonal of the matrix gives an idea about "resistance level" to a particular topic. In general, there is a prevalence of the grey color in the diagonal (17 values) which suggest that the clients are mostly open to continuing the same topic. With twelve black values and six white values, the diagonal also suggests that counselors tend to switch topics twice as the patients do. A possible explanation is that counselors aim at searching for other correlated symptoms or subjects that would lead to a mental disease. The other values of the matrix describe the second and third type of topic changes; the number of white and black values are approximately equal, which means that the conversations, in general, are discussed without perceived tactics. Nevertheless, some rows and columns are mostly either negative or positive (e.g. *Parenting*) which indicate the potential use of some strategies. The counselor often switches the topic if the previous one was *Mania*, *Medication* or *Patient-Counselor Relations*. Instead, he/she frequently starts a new topic if the patient's talk restrains less semantic contents (*Meaningless Talk*). On the other hand, the patient often switches topics if the previous topic was related to *Parenting*, *Friendship*, *Sexual dysfunction*, *Crying*, or *Stress-and-Work*.

6 Evaluation

The evaluation of the performance of a topic model is not an easy task. In most cases, topics need to be manually evaluated by humans, which may express different opinions and annotations. The most common quantitative way to assess

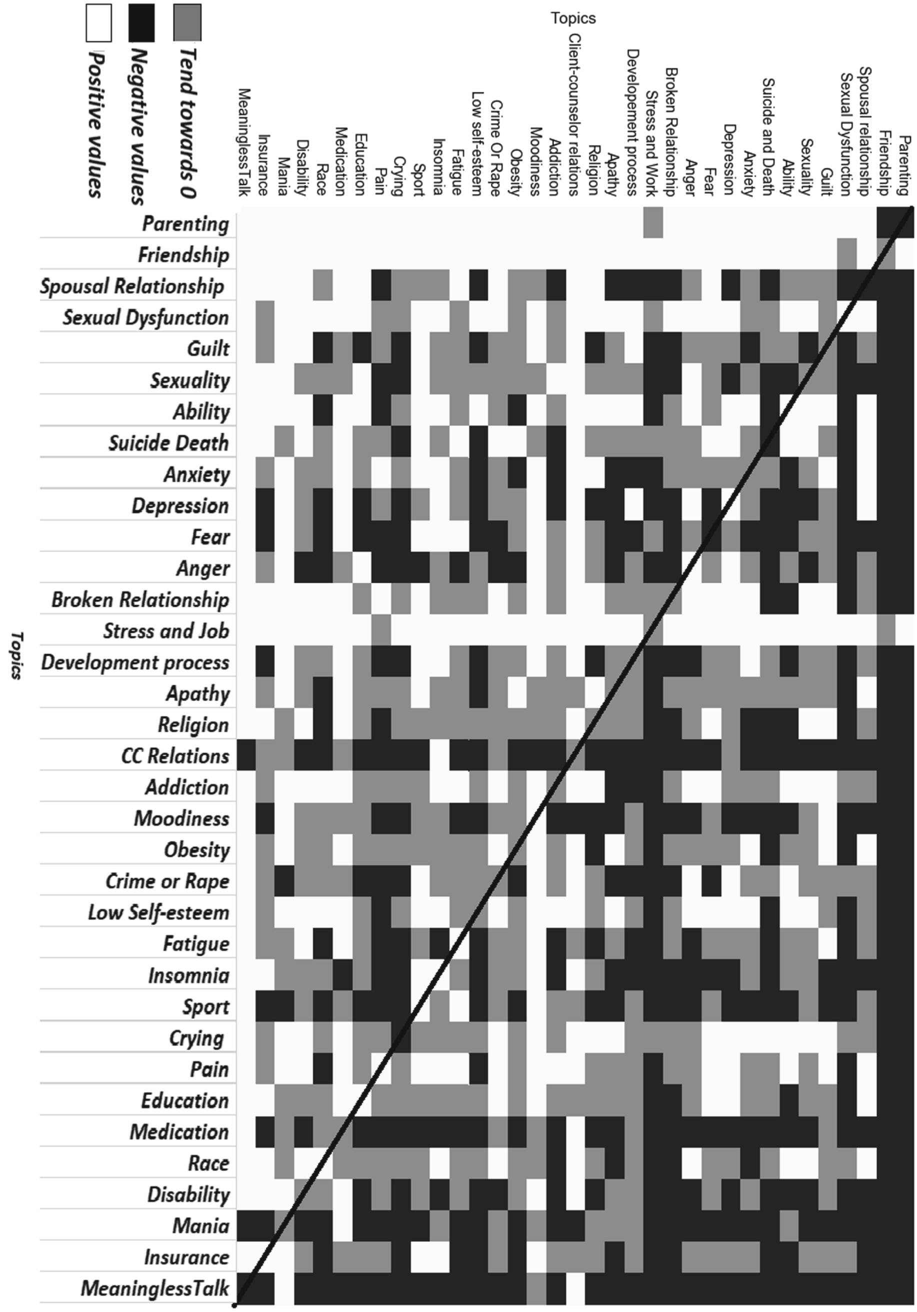

Fig. 6. The difference matrix between CP and PC.

a probabilistic model is to measure the log-likelihood of a held-out test set performing perplexity. However, the authors in [10] have shown that, surprisingly, perplexity and human judgment are often not correlated, and may infer less semantically meaningful topics. A potential solution to this problem is provided by the topic coherences, that is a typical way to assess qualitatively topic models by examining the most likely words in each topic. For such a purpose, we employed Palmetto[10], a tool to compute topic coherence of a given word set with six different methods. The one that we selected for our purposes was the C_V method [22], which uses word co-occurrences from the English Wikipedia, and that has been proven to highly correlate with human ratings. C_V is depended on a one-set segmentation of the top words and a measure that uses normalized pointwise mutual information. The one-set segmentation computes the cosine similarity between each top words vector and the amount of all top words vectors. The coherence value is then the arithmetic average of these similarities and represents an intuitive measure of the goodness of the topics produced by PLDA. In this work, we evaluated our PLDA topic model for topic detection using C_V coherence. In particular, we gave the top five terms (according to the weight of PLDA shown in Table 1) for each of the 34 topics as the input, obtaining as output a satisfactory coherence amongst all the detected topics. Indeed on average a topics coherence value larger than 50% was obtained, which is recognized in the research community already as a well-acceptable coherence score for a TDT model. This further substantiates the validity and potentials of our method.

7 Conclusions

In this paper, we study the topic propagation in a large collection of transcriptions from real psychotherapeutic sessions. For topic modeling, we used Partially Labeled Latent Dirichlet Allocation, PLDA, which makes it possible to track both common topics that are known in advance, and topics encountered in the conversational data. Moreover, we used topic coherence evaluation algorithms to evaluate the consistency of the topic system. Finally, we computed TTM to capture the dynamics of each ongoing topic in the conversations understanding the patterns how the patient and therapist, respectively, maintain and switch topics during the therapy sessions.

Knowing how topics change and propagate over the session can be used by counselors to drive the discussion and to adjust their assessment of the emotional state and barriers of the patient. These aspects of interaction are critical for all mental health specialists as they are related to the health state of the patients. We conclude that PLDA and TTM may be of benefit to the therapeutic conversational speech analysis and other real-life applications of AI to psychotherapy.

[10] http://aksw.org/Projects/Palmetto.html.

References

1. Allahyari, M., Kochut, K.: Automatic topic labeling using ontology-based topic models. In: 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA), pp. 259–264, December 2015
2. Angus, D., Smith, A.E., Wiles, J.: Human communication as coupled time series: quantifying multi-participant recurrence. IEEE Trans. Audio Speech Lang. Process. **20**(6), 1795–1807 (2012)
3. Angus, D., Watson, B., Smith, A., Gallois, C., Wiles, J.: Visualising conversation structure across time: insights into effective doctor-patient consultations. PloS ONE **7**(6), e38014 (2012)
4. Asuncion, A., Welling, M., Smyth, P., Teh, Y.W.: On smoothing and inference for topic models. In: UAI 2009, pp. 27–34. AUAI Press, Arlington (2009)
5. Atkins, D.C., Steyvers, M., Imel, Z.E., Smyth, P.: Scaling up the evaluation of psychotherapy: evaluating motivational interviewing fidelity via statistical text classification. Implementation Sci.: IS **9**, 49 (2014)
6. Bangalore, S., Di Fabbrizio, G., Stent, A.: Learning the structure of task-driven human-human dialogs. IEEE Trans. Audio Speech Lang. Process. **16**(7), 1249–1259 (2008)
7. Blei, D.M., Lafferty, J.D.: Dynamic topic models. In: ICML 2006, pp. 113–120. ACM, New York (2006)
8. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet allocation. J. Mach. Learn. Res. **3**, 993–1022 (2003)
9. Breuing, A., Wachsmuth, I.: Talking topically to artificial dialog partners: emulating humanlike topic awareness in a virtual agent. In: Filipe, J., Fred, A. (eds.) ICAART 2012. CCIS, vol. 358, pp. 392–406. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-36907-0_26
10. Chang, J., Boyd-Graber, J., Gerrish, S., Wang, C., Blei, D.M.: Reading tea leaves: how humans interpret topic models. In: NIPS 2009, pp. 288–296. Curran Associates Inc., Red Hook (2009)
11. Chen, W.T., Lin, S.C., Huang, S.L., Chung, Y.S., Chen, K.J.: E-HowNet and automatic construction of a lexical ontology. In: COLING 2010, pp. 45–48. Association for Computational Linguistics, Stroudsburg (2010)
12. Chen, Y., Liu, L.: Development and research of topic detection and tracking. In: 2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS), pp. 170–173, August 2016
13. Drew, P., Chatwin, J., Collins, S.: Conversation analysis: a method for research into interactions between patients and health-care professionals. Health Expect.: Int. J. Publ. Particip. Health Care Health Policy **4**(1), 58–70 (2001)
14. Gaut, G., Steyvers, M., Imel, Z.E., Atkins, D.C., Smyth, P.: Content coding of psychotherapy transcripts using labeled topic models. IEEE J. Biomed. Health Inform. **21**(2), 476–487 (2017)
15. Gelbukh, A.: Natural language processing. In: Fifth International Conference on Hybrid Intelligent Systems (HIS 2005), pp. 1–6 (2005)
16. Howes, C., Purver, M., McCabe, R.: Investigating topic modelling for therapy dialogue analysis. In: Proceedings of the IWCS 2013 Workshop on Computational Semantics in Clinical Text (CSCT 2013), pp. 7–16. Association for Computational Linguistics (2013)
17. Imel, Z.E., Steyvers, M., Atkins, D.C.: Computational psychotherapy research: scaling up the evaluation of patient-provider interactions. Psychotherapy (Chicago, Ill.) **52**(1), 19–30 (2015)

18. Mohr, J.W., Bogdanov, P.: Introduction-topic models: what they are and why they matter. Poetics **41**(6), 545–569 (2013). Topic Models and the Cultural Sciences
19. Khin, N.P.P., Aung, T.N.: Analyzing tagging accuracy of part-of-speech taggers. In: Zin, T.T., Lin, J.C.-W., Pan, J.-S., Tin, P., Yokota, M. (eds.) GEC 2015. AISC, vol. 388, pp. 347–354. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-23207-2_35
20. Ramage, D., Hall, D., Nallapati, R., Manning, C.D.: Labeled LDA: a supervised topic model for credit attribution in multi-labeled corpora. In: EMNLP 2009, pp. 248–256. Association for Computational Linguistics, Stroudsburg (2009)
21. Ramage, D., Manning, C.D., Dumais, S.: Partially labeled topic models for interpretable text mining. In: KDD 2011, pp. 457–465. ACM, New York (2011)
22. Röder, M., Both, A., Hinneburg, A.: Exploring the space of topic coherence measures. In: WSDM 2015, pp. 399–408. ACM, New York (2015)
23. Uysal, A.K., Gunal, S.: The impact of preprocessing on text classification. Inf. Process. Manag. **50**(1), 104–112 (2014)
24. Yeh, J.F., Tan, Y.S., Lee, C.H.: Topic detection and tracking for conversational content by using conceptual dynamic latent Dirichlet allocation. Neurocomputing **216**(Suppl. C), 310–318 (2016)