An Unsupervised Approach to Extract Life-Events from Personal Narratives in the Mental Health Domain

Seyed Mahed Mousavi, Roberto Negro, Giuseppe Riccardi
Signals and Interactive Systems Lab, University of Trento, Italy
{mahed.mousavi, giuseppe.riccardi} @unitn.it

Abstract

Personal Narratives are an important source of knowledge in the mental health domain. Over an extended period of time, the psychologist learns about the patient's life-events and participants from the Personal Narratives shared during each therapy session. The acquired knowledge is then used to support the patient to reach a healthier mental state by appropriate targeted feedback during each conversation. In this work, we propose an unsupervised approach to automatically extract personal life-events and participants from the patient's narratives and represent them as a personal graph. This personal graph is then updated at each interaction with the patient. We have evaluated our proposed approach on a dataset of longitudinal Italian Personal Narratives as well as a dataset of English commonsense stories.

1 Introduction

There is a growing research and clinical interest in developing conversational agents (CA) for mental health support as Personal Healthcare Agents (PHA) (Abd-alrazaq et al., 2019; Fitzpatrick et al., 2017; Inkster et al., 2018). However, the lack of appropriate domain knowledge has resulted in the abundance of rule-based dialogue systems in the mental health domain with shallow interactions and weak user engagement (Abd-Alrazaq et al., 2021). Currently available dialogue knowledge can be adequate for consumer-oriented agents or holding a free-topic social conversation. However, it can not be used to hold a dialogue about personal life-events and emotions. Meanwhile, patients' conversations in the mental health domain have a unique and complex structure since they encompass personal feelings and situations which vary across patients and interventions.

In order to carry out a personal conversation regarding the patient's life-events, it is essential to obtain the required knowledge during each interaction with the patient and from her Personal Narratives. Personal Narratives (PN) are recollections of thoughts and emotions about life-events of the patient. These narratives are used by the psychologist to identify the issues that have activated the patient's emotional state and provide support accordingly in order to reach a healthier mental status (Tammewar et al., 2019; Vromans and Schweitzer, 2011).

In this work, we present an unsupervised approach, inspired by (Chambers and Jurafsky, 2008), to automatically extract the life-events and their participants from the patient's PNs, and construct a Personal Space Graph. Figure 1 represents the work flow of our model. Through the interaction with the patient, each narrative is parsed and presented in terms of its predicates (the events, the edges of the graph) and their noun dependencies (the participants, the nodes of the graph). Each edge has an index based on its order of appearance in the narrative which makes it possible to reconstruct the order of occurrences among the events (for instance, the event "*litigo spesso*" is mentioned after "*parla male*"). Moreover, the events and participants mentioned in a recent narrative are considered to be more relevant for an ongoing interaction. Based on this assumption, older nodes and edges in the graph will become less relevant upon receiving a new narrative (presented by dashed lines in Figure 1). The obtained graph can be integrated with PHAs to automatically identify the life-event that is distressing the patient from his/her PNs to provide support and monitor its recurrence.

We have evaluated our approach on a dataset

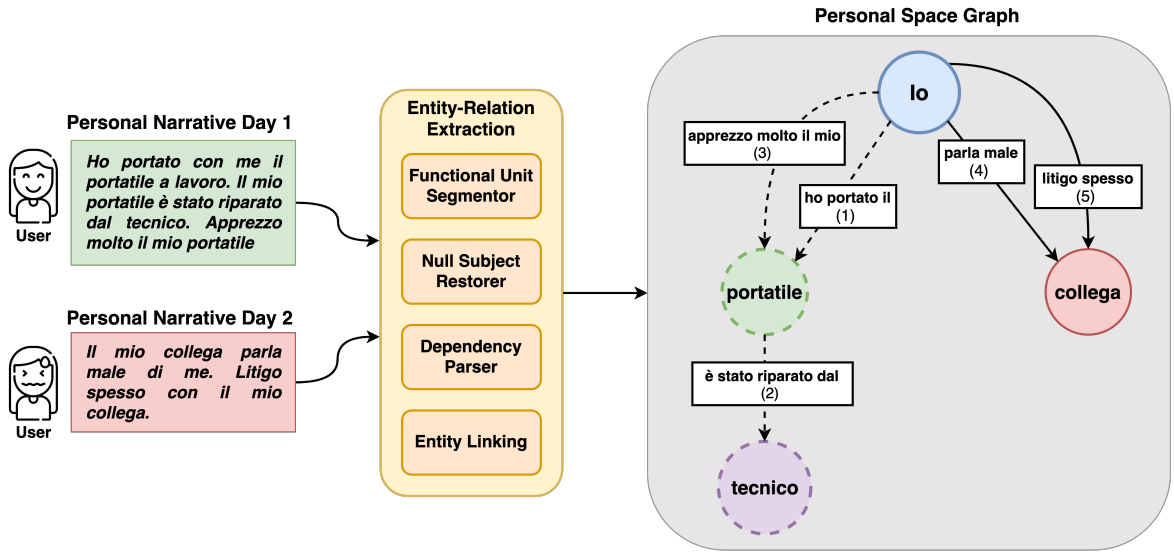

Figure 1: Each Personal Narrative (PN) is presented as a graph of patient's personal space of participants and events. Each edge and the adjacent nodes stand for an event and its corresponding participants in the narrative, respectively. Each edge (event), has an index based on its appearances in the PN which makes it possible to reconstruct the order of occurrences among the events. Moreover, events and participants extracted from prior narratives are considered less relevant for an ongoing interaction and have a lower importance score, presented by dashed lines.

of longitudinal Italian PNs collected from patients who were receiving Cognitive Behavioural Therapy to manage their distress[i]. Besides, the English adaptation of our model was evaluated in the "Story Cloze Test" setting introduced by (Mostafazadeh et al., 2016). The results show that the proposed approach obtains similar results to other unsupervised models on the English dataset, and can be a strong baseline for personal space representation and response selection in Italian.

2 Literature Review

Unsupervised Event Extraction There have been several interesting works regarding the unsupervised extraction of events and their participants from unstructured text. (Chambers and Jurafsky, 2008) introduced the concept of "Narrative Event Chain". In this work, the events with a shared participant are assumed to be parts of a uniform story. They present the events in the narrative by the verbs that have a shared participant, and the participant's role for each verb. (Chambers and Jurafsky, 2009) then extended this model to "Narrative Schema", obtained by merging different event chains extracted from one narrative into an integrated uniform schema in order to model the document by all participants across the verbs. Recently, (Hatzel and Biemann, 2021) proposed to further extend the "Narrative Schema" concept to support long documents in German language by 1) performing language adaptation of the model; and 2) dividing the event sequence into multiple strongly-connected schema in order to present different scenes in a long story.

Evaluation Criteria Regarding the evaluation of the models, the mentioned unsupervised approaches were evaluated in "Narrative Cloze Test" setting. In this setting, an event is removed from a sequence of events in a document and the task is to predict the most probable candidate for the missing event (Chambers and Jurafsky, 2008). Later however, (Mostafazadeh et al., 2016) introduced "Story Cloze Test" evaluation criterion. In this setting, the system selects a complete lexicalized sentence as the closure to a story rather than predicting the missing event. For this purpose, the authors crowd-sourced a dataset of commonsense stories, called ROCStories, with right and wrong ending sentences for each story.

3 Personal Space Graph Representation

In this work, we propose an unsupervised Entity-Relation Extraction (ERE) approach to obtain the

[i]This data collection has been approved by the Ethical Committee of the University of Trento

personal graph of life-events and participants from the user's PNs in the mental health domain. Figure 1 shows the workflow of our approach, consisting of five main modules.

Functional Unit Segmentor Upon receiving a narrative, it is first segmented into its functional units. A functional unit is a contiguous span within a message which has a coherent communicative intention (Oltean et al., 2017). The segmentation into Functional Units was performed by a seq2seq model (Zhao and Kawahara, 2019), trained to jointly perform Functional Unit segmentation and Dialogue Act (DA) tagging, based on ISO standard DA tagging in Italian (Roccabruna et al., 2020). The model was trained on the corpus of Italian dialogues in the mental health (Mousavi et al., 2021). The predictions of the model were then post-edited and adjusted by two human annotators with strong inter-annotator agreement (0.87) measured by Cohen's κ coefficient (Fournier and Inkpen, 2012).

Dependency Parser Each functional unit is then passed to the dependency parser to obtain the corresponding dependency tree, for which spaCy natural language processing library[ii] was used. Using the obtained tree and part-of-speech tags, tokens tagged as nouns and proper nouns are extracted as nodes in the graph (nominal modifier nouns are excluded in this process since they are describing/specifying characteristics of another noun). In cases that pronouns are subjects or objects of a verb, they are extracted as nodes as well.

Entity Linking In order to make sure repeated nouns or variations of the same noun are mapped to the correct node in the graph, an Entity Linking module is defined. This module queries BabelNet[iii] and ConceptNet[iv] semantic networks for the root form of the extracted nouns and matches them consequently to obtain a set of entities and participants in the narrative.

Null Subject Restorer All the verbs contained in the functional unit are extracted and controlled for possible null subject case. Null subjects are non overtly expressed subject pronouns commonly used in pro-drop languages such as Italian and Spanish (Russo et al., 2012). In this case, the subject of the verb is restored as a pronoun based on its conjugation using an out-of-the-shelf library MLCONJUG3[v] to make sure each event participant is detected and extracted correctly.

Entity-Relation Extraction Lastly, the model navigates through the dependency tree to find the verbs that connect the extracted entities as subjects and objects/oblique nominals. In cases of entity conjunctions, the same verb spans over all the entities in the same conjunction. For a better visualization, the neighbours of the verb in the dependency tree are explored to obtain an entire predicate composed by adverbs, ad-positions and auxiliaries as the edge of the graph.

The obtained graph is specific to each patient and spans over the life-events shared in the narratives. In each graph, the patient is presented as the node "*Io*" and all the other participants in the patient's PNs are connected to it by the corresponding predicate. PNs in the mental health domain are about the events that activated the patient's emotional state. Therefore, it is important to maintain the consecutive order among events in each PN as well as among subsequent PNs through several interactions with the patient. For this purpose, each edge is indexed based on its sequence of appearance in the narrative in order to reconstruct the ordered chain of events. Moreover, events extracted from prior narratives are considered less relevant to the patient's mental status, unless they re-appear in recent narratives. Therefore, these events receive lower importance score in time based on the assumption that the issue is resolved and the patient does not feel the need to re-mention it.

4 Evaluations

We have evaluated our proposed approach in two different settings in the mental health domain for Italian language. Furthermore, we have compared the performance of its English adaptation with other models in the "Story Cloze Test" setting introduced by (Mostafazadeh et al., 2016).

4.1 Personal Narratives Evaluation

We first collected a dataset PNs from Italian patients who were receiving Cognitive Behavioural Therapy to better manage their distress[vi]. Using the approach introduced priorly by (Mousavi et al., 2021), the patients were asked to write PNs

[ii] spaCy `spacy.io`

[iii] BabelNet `babelnet.org`

[iv] ConceptNet `conceptnet.io`

[v] MLCONJUG3 `pypi.org/project/mlconjug3`

[vi] This data collection has been approved by the Ethical Committee of the University of Trento

Closure Selection (Pool of 2 & 5)					Narrative Selection (Pool of 2)				
Recall	**Rand.**	**TF-IDF**	**Nouns**	**ERE**	**History Size**	**Rand.**	**TF-IDF**	**Nouns**	**ERE**
R@1 in 2	50%	71.1%	41.3%	59.0%	**2 Personal Narratives**	50%	74.4%	68.8%	71.4%
R@1 in 5	20%	51.6%	34.8%	42.7%	**5 Personal Narratives**	50%	75.3%	68.8%	72.0%

Table 1: The results of evaluating our model for Entity Relation Extraction in Italian (ERE) in two different selection settings at closure level and narrative level on a dataset of Personal Narratives collected from patients in the mental health domain.

	NC-AP	**NC-ROC**	**Nouns**	**ERE**
R@1 in 2	48.7	49.4	45.1	45.6

Table 2: The result of evaluating the English adaptation of our model in "Story Cloze Test" setting, compared with other unsupervised approaches (Mostafazadeh et al., 2016). NC-AP and NC-ROC models stand for the standard Narrative Event Chain model (Chambers and Jurafsky, 2008), with the point-wise mutual information function train on Associate Press (AP) portion of the English Gigaword Corpus and the ROCStories, respectively.

about real-life situations and events that have activated their emotional state for the period of three months. As the result, we collected 241 PNs from 18 patients with average length of 128.2 tokens per PN and average number of 11.9 PNs per patient.

Using the obtained dataset of PNs, in the first setting we evaluated the model for the task of Closure Selection. That is, the model was tasked to select the correct closure sentence for an incomplete narrative based on the participants and events (verbs) it consists of. Similar to a response-selection setting, we assessed the performance of the model using two pools of 2 and 5 candidates, each consisting of 1 correct closure and *n-1* distractors.

In the second setting, we evaluated whether the obtained graph can correctly represent a personal space of events and participants that varies for each user. To this end, the model was first presented with a set of 2 or 5 consecutive PNs from a specific patient as history. Once the corresponding personal space graph was extracted, the model was tasked to select the next possible PN from that patient from a pool of 2 candidates, consisting of the correct PN and a distractor (a PN written by a different user.)

The results of these evaluations are presented in Table 1. In the first scenario, while TF-IDF manages to be a strong baseline, our proposed system outperforms the Random baseline and has a much higher success rate than the selection solely based on the recurrence of the nouns. Moreover, by raising the task difficulty and increasing the pool size to 5, our model maintains the same performance trend. Regarding the second evaluation, the results indicate that the recurrence of the nouns is an important factor for the model to select the next possible PN. Nevertheless, our model manages to outperform this baseline by considering the predicates as an additional factor, and get closer to TF-IDF scores.

4.2 Story Cloze Test

In order to compare the performance of our model with other unsupervised approaches, the English adaption of the model was evaluated in the "Story Cloze Test" setting. In this setting, the model is tasked to select the most probable ending for a four-sentence story from a pool of 2, consisting of the right ending and the wrong one. (Mostafazadeh et al., 2016). The result of this evaluation for the test set of 3744 stories is presented in Table 2, indicating that our model performance is inline with other unsupervised approaches.

5 Conclusion

In this work, we present an approach to automatically extract life-events and participants from patients' Personal Narratives in the mental health domain and represent them as a personal graph. This graph can be a source of knowledge for Personal Healthcare Agents (PHA) in this domain, to automatically identify the life-event that is activating the user's emotional state and causing distress.

We evaluated our model on a domain-specific dataset of Personal Narratives in Italian as well as an open-domain dataset of commonsense stories in English. The results indicate that our proposed model performs in-line with other unsuper-

vised alternatives and can be a strong baseline for automatic extraction of life-events from Personal Narratives in Italian.

Acknowledgments

The research leading to these results has received funding from the European Union – H2020 Programme under grant agreement 826266: COADAPT.

References

Alaa A Abd-alrazaq, Mohannad Alajlani, Ali Abdallah Alalwan, Bridgette M Bewick, Peter Gardner, and Mowafa Househ. 2019. An overview of the features of chatbots in mental health: A scoping review. *International Journal of Medical Informatics*, 132:103978.

Alaa A Abd-Alrazaq, Mohannad Alajlani, Nashva Ali, Kerstin Denecke, Bridgette M Bewick, and Mowafa Househ. 2021. Perceptions and opinions of patients about mental health chatbots: Scoping review. *Journal of Medical Internet Research*, 23(1):e17828.

Nathanael Chambers and Dan Jurafsky. 2008. Unsupervised learning of narrative event chains. In *Proceedings of ACL-08: HLT*, pages 789–797.

Nathanael Chambers and Dan Jurafsky. 2009. Unsupervised learning of narrative schemas and their participants. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, pages 602–610.

Kathleen Kara Fitzpatrick, Alison Darcy, and Molly Vierhile. 2017. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): a randomized controlled trial. *JMIR mental health*, 4(2):e19.

Chris Fournier and Diana Inkpen. 2012. Segmentation similarity and agreement. *arXiv preprint arXiv:1204.2847*.

Hans Ole Hatzel and Chris Biemann. 2021. Towards layered events and schema representations in long documents. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Student Research Workshop*, pages 32–39, Online, June. Association for Computational Linguistics.

Becky Inkster, Shubhankar Sarda, and Vinod Subramanian. 2018. An empathy-driven, conversational artificial intelligence agent (wysa) for digital mental well-being: real-world data evaluation mixed-methods study. *JMIR mHealth and uHealth*, 6(11):e12106.

Nasrin Mostafazadeh, Nathanael Chambers, Xiaodong He, Devi Parikh, Dhruv Batra, Lucy Vanderwende, Pushmeet Kohli, and James Allen. 2016. A corpus and cloze evaluation for deeper understanding of commonsense stories. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 839–849.

Seyed Mahed Mousavi, Alessandra Cervone, Morena Danieli, and Giuseppe Riccardi. 2021. Would you like to tell me more? generating a corpus of psychotherapy dialogues. In *Proceedings of the Second Workshop on Natural Language Processing for Medical Conversations*, pages 1–9, Online, June. Association for Computational Linguistics.

Horea-Radu Oltean, Philip Hyland, Frédérique Vallières, and Daniel Ovidiu David. 2017. An empirical assessment of rebt models of psychopathology and psychological health in the prediction of anxiety and depression symptoms. *Behavioural and cognitive psychotherapy*, 45(6):600–615.

Gabriel Roccabruna, Alessandra Cervone, and Giuseppe Riccardi. 2020. Multifunctional iso standard dialogue act tagging in italian. *Seventh Italian Conference on Computational Linguistics (CLiC-it)*.

Lorenza Russo, Sharid Loáiciga, and Asheesh Gulati. 2012. Improving machine translation of null subjects in italian and spanish. In *Proceedings of the Student Research Workshop at the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pages 81–89.

Aniruddha Tammewar, Alessandra Cervone, Eva-Maria Messner, and Giuseppe Riccardi. 2019. Modeling user context for valence prediction from narratives. In *Interspeech*, page 3252–3256.

Lynette P Vromans and Robert D Schweitzer. 2011. Narrative therapy for adults with major depressive disorder: Improved symptom and interpersonal outcomes. *Psychotherapy research*, 21(1):4–15.

Tianyu Zhao and Tatsuya Kawahara. 2019. Joint dialog act segmentation and recognition in human conversations using attention to dialog context. *Computer Speech & Language*, 57:108–127.