
Article

Not peer-reviewed version

Leveraging Medical Discourse to Answer Complex Questions

[Boris Galitsky](#) *

Posted Date: 29 December 2023

doi: [10.20944/preprints202312.2149.v1](https://doi.org/10.20944/preprints202312.2149.v1)

Keywords: Question answering system in health; computational medical discourse; large language models



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Leveraging Medical Discourse to Answer Complex Questions

Boris Galitsky

bgalitsky@hotmail.com

Abstract: We review the literature on medical discourse and attempt to build a computational model of it. Medical discourse sheds a light on communication structure of patient-doctor and other communication scenarios in healthcare and should be leveraged to facilitate and automate this communication when it is possible and practical. We propose a unified framework to represent communication discourse at the meta-level, where the subject of the communication is expressed in a language object. So far, the broad range of work on medical discourse is detached from computational discourse analysis, and we explore the possibilities of filling this gap and computationally treat the peculiarities of how information is passed between the agents in a hospital setting. We select the domain of question answering (QA) against a corpus of medical documents of diverse nature to evaluate our computational model of medical discourse. It turns out that applying specific structures obtained in medical discourse studies improves the relevance and efficiency of question answering.

Keywords: question answering system in health; computational medical discourse; large language models

1. Introduction to Social Discourse

The concepts of text and discourse have encountered a multitude of interpretations within the social science community. Virtually every paper or article delves into these notions, often citing influential figures such as Michel Foucault. Consequently, the term "discourse" encompasses a vast array of meanings, ranging from historical monuments, policies, and political strategies to narratives, texts, speeches, and even broader language-related discussions. The term has expanded to include racist discourse, gendered discourse, discussions on employment, media discourse, populist discourse, historical discourses, and more. This broadening of the discourse concept extends its definition from a genre to a register or style and from a physical structure to a political agenda. This expansion has led to confusion, resulting in criticism and misunderstandings.

In this chapter, our aim is to formalize a comprehensive understanding of discourse structures in communication. Once successfully formalized, covering various examples in the literature corpus, we will transition to exploring question-answering (QA) applications within the healthcare domain. The primary focus will be on answering questions based on a collection of medical documents. In terms of a neuro-symbolic architecture, a Language Model (LLM) is complemented by a symbolic discourse model of a document. This integration embeds both language-specific elements and meta-language information into a neural representation for effective QA.

1.1. Metalinguage Model of Discourse

In the context of linguistics and philosophy of language, "language-object" and "metalinguage" refer to different levels of linguistic analysis. *Language-object* typically refers to the actual language being studied or discussed. It is the primary focus of linguistic analysis, representing the object of inquiry. For example, if linguists are studying English, then English itself is the language-object. The language-object is the system of symbols, rules, and structures that people use to communicate, and it is the subject of investigation in linguistics. *Language-object* includes syntax and semantics.

Metalanguage refers to a second-level language that is used to talk about or describe the language-object. It is a language used to discuss the features, structure, and elements of the primary language-object. In other words, it is a language that analyzes, interprets, or comments on the language being studied. The metalanguage provides a framework for discussing linguistic concepts and phenomena. Discourse describes a structure of how a text author organized her thoughts in text; therefore, discourse, unlike semantics, is expressed in metalanguage.

In summary, the language-object is the language under study, while the metalanguage is the language used to analyze and discuss the properties and structures of the language-object. This distinction is fundamental in linguistic analysis and will help us communicate about language discourse in a systematic and precise manner.

Discourse analysis provides a general framework to *problem-oriented social research*. The problem-oriental part is expressed in language-object, and social part (ascending from the level of an individual) – in metalanguage. It allows the integration of different dimensions of *interdisciplinarity* and multiple perspectives on the object investigated (some perspectives are expressed in language-object, other in metalanguage). *Discourse analysis* allows the integration of different dimensions of *interdisciplinarity* into the metalevel and multiple perspectives on the object investigated. Every interview, focus group debate, TV debate or visual symbol is conceived as a *semiotic* entity, embedded in an immediate, text-internal co-text and an intertextual and socio-political context. Analysis thus has to take into account the *intertextual and interdiscursive relationships* (see below) between utterances, texts, genres and discourses, as well as the extra-linguistic social/sociological variables, the *history and archaeology of an organization*, and institutional frames of a specific *context of situation*. Semiotic entity is assumed to be an object level information conveying entity in contrast to an entity about relations between object-level entities, which is meta-level and not semiotic, in our definition.

Semiotics is the study of signs and symbols and their use or interpretation. It is a field of study that examines how meaning is created and communicated through signs and symbols in various contexts. Semiotics is often associated with linguistics, but it extends beyond language to encompass a wide range of cultural and social phenomena. Semiotics analyzes signs and symbols, which can take various forms such as words, images, sounds, gestures, and objects. A sign is anything that communicates meaning. Different cultures and societies have their own semiotic systems, which are sets of rules and conventions governing the use and interpretation of signs and symbols. Semiotic Analysis deconstructs and understands the meanings embedded in signs and symbols. Semiotic analysis can be applied to various fields, including literature, art, film, advertising, and everyday communication.

Intertextuality is an important notion to define a graph structure of discourse. *Intertextuality* refers to the fact that all texts are linked to other texts; these links can be labeled by temporal relationships (Zengin 2016). Such links can be established in different ways: through continued reference to a topic or main actors; through reference to the same events; or by the transfer of main arguments from one text into the next. The latter process is also labeled recontextualization. De-contextualization occurs when a linguistic element, such as a word or phrase, is considered independently of the surrounding linguistic context. This process is often necessary for analyzing the inherent meaning of the linguistic unit itself. For example, de-contextualization is needed for fact-checking (Chap ??).

By taking an argument and restating it in a new context, we first observe the process of *de-contextualization*, and then, when the respective element is implemented in a new context, of *recontextualization*. The element then acquires a new meaning because meanings are formed in use (Wittgenstein 1967). Moreover, these meaning formation processes occur under metalevel control. *Interdiscursivity*, on the other hand, indicates that discourses form a hierarchical structure and are linked to each other in various ways: discourse of a paragraph, a section, an utterance in a dialogue, the whole dialogue and the whole document. If we define discourse as primarily topic-related, i.e. a discourse on X, then a discourse on 'unemployment' often refers, for example, to topics or subtopics of other discourses, such as gender or racism: arguments on systematically lower salaries for women or migrants might be included in discourses on employment (see below for definitions of text and

discourse). In this case, discourse is a meta-theory for the theory expressed by a set of documents related to employment.

According to Fairclough (2003), discourse is characterized as the utilization of language viewed as a manifestation of social practice, establishing a dialectical relationship between the microstructure of discourse (including linguistic structures like various interactional features) and the macrostructures of society, encompassing social structure, ideology, and power. In this context, the social level functions as a meta-level concerning an object level, such as communication in the field of health.

1.2. Semantic Network and Discourse

In the Introduction, we look at the discourse from the broad perspective and focus on medical discourse in Sect. 2, 6 and 7.

Effective collaborative discourse in a healthcare organization necessitates the active involvement of all parties, encompassing both cognitive and social engagement. To explore the intricate socio-cognitive dynamics inherent in collaborative discourse, (Chen et al., 2022) suggests modeling it as a socio-semantic network and subsequently utilizing network motifs, defined as recurring, meaningful subgraphs, to characterize both the network and the discourse itself.

Since the emergence of the socio-cognitive paradigm four decades ago, the significance of social interaction has been underscored in various educational theories and practices. For instance, within a team learning framework, members collaboratively learn by constructing a shared problem space, building upon each other's contributions, and creating knowledge artifacts collectively. Collaborative discourse draws inspiration from the socio-cognitive paradigm of learning, where participants engage in substantive discussions related to a particular domain. By utilizing the interpersonal and intersubjective space, learners are expected to comprehend new concepts and collectively construct shared knowledge that extends beyond the grasp of individual understanding.

Collaborative discourse represents an effort to harness the interpersonal communication and intersubjective meaning-making of decision-makers and learners, aiming to achieve learning objectives beyond individual capabilities (Stahl and Hakkarainen, 2021). Rooted in socio-cognitive perspectives of learning, collaborative discourse seeks to leverage both cognitive and social processes, encouraging learners to engage in activities like articulation, explanation, questioning, and collaborative knowledge co-construction. In contrast to passive and active learning, constructive conditions involve utilizing prior knowledge to interpret information, while interactive conditions entail collaboratively co-constructing solutions or elaborating on each other's ideas (Chi and Wylie, 2014). The intricate dynamics of social and cognitive processes characterize sophisticated collaborative discourse in advanced interactive conditions.

The socio-semantic motif framework operates under the fundamental assumption that collaborative discourse necessitates multiple collaborators engaged in discussions about shared content (Figure 1). Both social interaction and shared attention are pivotal components of collaborative discourse. Without meaningful interactions, discourse cannot truly be collaborative, and if a group merely shares content without semantic overlap, the intersubjective meaning-making crucial for collaboration is unlikely to occur.

In this framework, socio-semantic network motifs are the basic building blocks, consisting of minimal sets of social and semantic entities. Each socio-semantic network motif in our framework comprises two learners and two words, forming six potential links (see Figure 1). The naming system, based on the number of edges on each layer, follows the proposal by (Chen et al., 2022). For example, $A(1,0)$ has one edge on the top layer, while $C(1,1)$ adds another edge between two layers. $A(0,2a)$ and $A(0,2b)$ share the same number of between-layer edges but differ in edge combinations.

In the domain of network science, network motifs have found extensive application in analyzing diverse networks, spanning biological, technological, infrastructural, and social domains (Milo et al., 2004). Within the field of environmental science, the focus shifts to two-layer network motifs, particularly relevant for probing socio-ecological systems that intertwine social actors (e.g., trip

organizers and airplane pilots) and ecological resources (e.g., mountains, airspace). Extending this framework to the domain of collaborative discourse, it employs two-layer network motifs.

In the investigation of socio-semantic networks, this framework aims to characterize discourse by identifying and assessing the frequency and significance of socio-semantic network motifs. Extracting all potential motifs from a socio-semantic network, the framework seeks to offer insights into discourse dynamics. The frequency of these motifs, along with their significance when compared to null models, is anticipated to furnish valuable indicators elucidating the nature of discourse as a socio-semantic system.

Chen et al. (2022) crafted two-layer socio-semantic networks, where the upper layer portrays the undirected interaction network among students, while the lower layer comprises high-frequency words derived from students' written discourse over a specific week. The links between a student and a word signify that the word was mentioned at least twice in the student's posts related to a particular reading. However, links between words themselves were not taken into account. An exemplar socio-semantic network, generated from discourse centered around a specific reading, is illustrated in Figure 2.

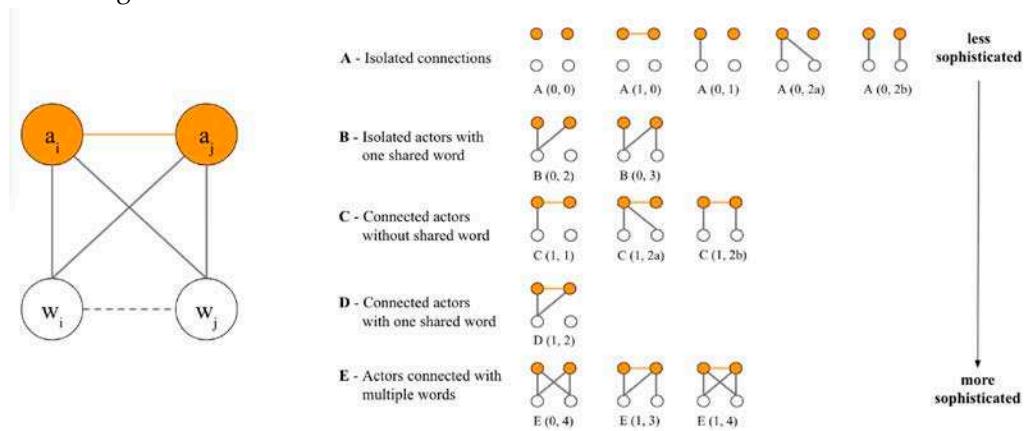


Figure 1. A socio-semantic network motif, and a classification of motifs in collaborative discourse.

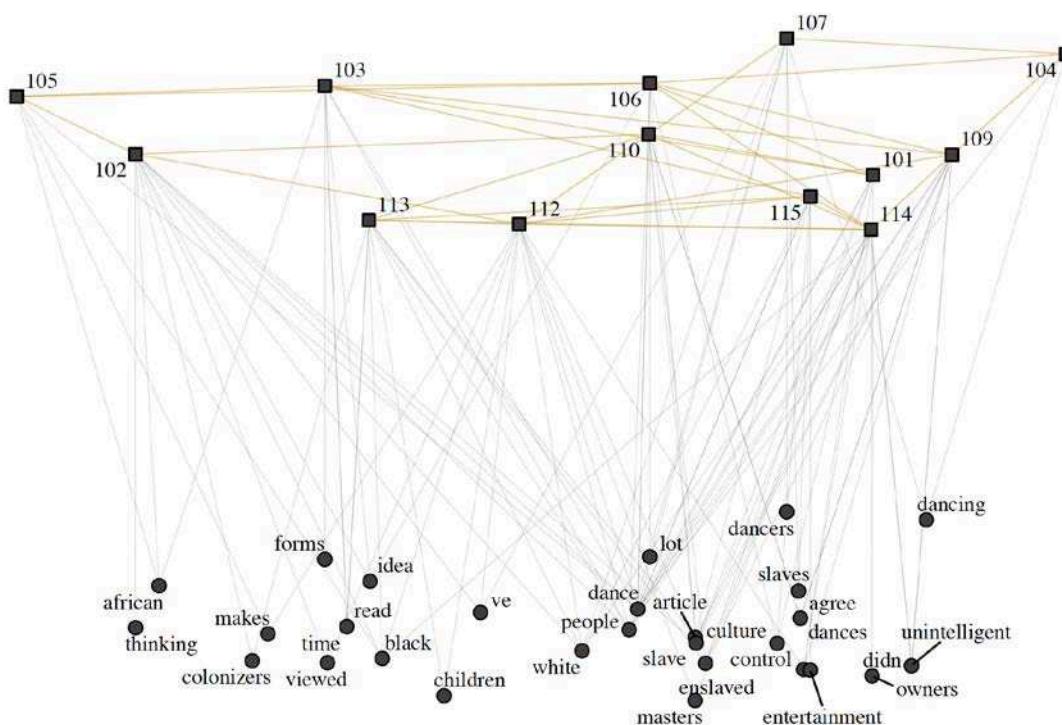


Figure 2. A socio-semantic network created from discourse around a particular reading.

A semantic network serves as a knowledge structure that visualizes the relationships among concepts, leveraging AI programming to extract data, establish connections between concepts, and highlight relationships. These networks capture the thematic connections discerned in focus group discussions. As we constructed networks based on macro-topics discussed and their interrelations, we systematically analyzed transcripts, identifying various topics and their argumentative development. For each relevant and significant topic identified, (Reisigl et al., 2009) introduced a new node, illustrating lines that represent the discursive connections (links) between existing topics. It is also possible to define additional relationships if needed. Such a meta-representation diagram for a focus group empowers researchers to formulate initial hypotheses regarding interaction dynamics and the flow of arguments. Figure 3 presents a semantic network for a group engaged in discussing security issues and Austrian neutrality.

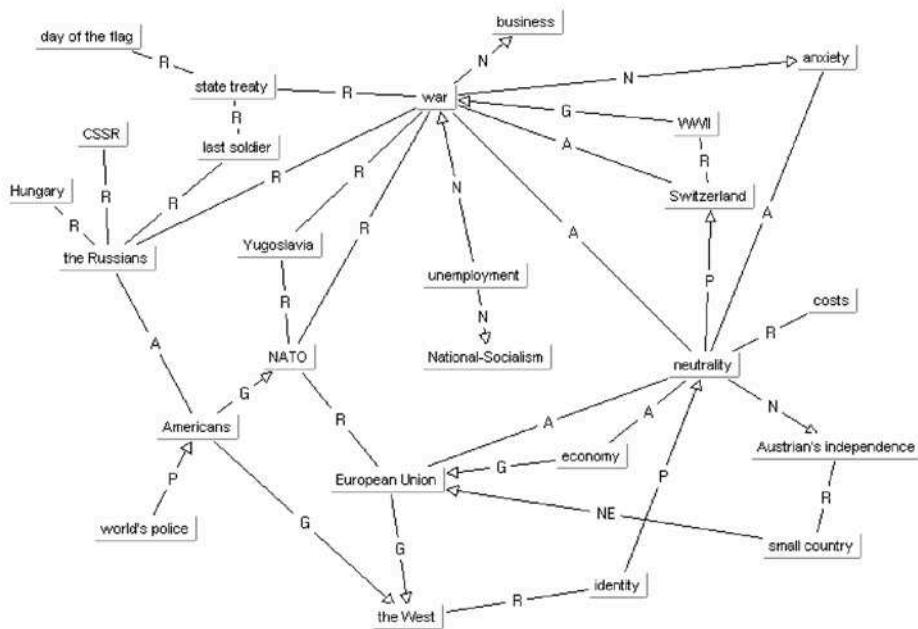


Figure 3. Semantic network assisting in discourse analysis of a discussion group.

1.3. Discourse in a Broader Sense

In a concise overview of social discourse domain, Van Dijk (2009) encapsulates the history of discourse studies (and underscores that the fundamental essence of this emerging discipline is the systematic and explicit analysis of diverse structures and strategies inherent in various levels of text and discourse. Consequently, discourse studies necessitates drawing upon a spectrum of disciplines including *anthropology, history, rhetoric, stylistics, conversation analysis, literary studies, cultural studies, pragmatics, philosophy, sociolinguistics*, and more (Reisigl et al. 2009).

One of the most salient features of the *discourse-historical approach* for example, is its endeavor to work interdisciplinarily, multi-methodically and on the basis of a variety of different empirical data. Depending on the object of investigation, it attempts to transcend the purely linguistic dimension and to include more or less systematically the historical, political, sociological and/or psychological dimensions in the analysis and interpretation of a specific discursive event. Thus, the triangulatory approach is based on a concept of *context* which takes into account four levels illustrated in Figure 4:

- (1) the immediate, language or text internal co-text (object-level)
- (2) the intertextual and interdiscursive relationship between utterances, texts, genres and discourses (meta-level)
- (3) the extra-linguistic social/sociological variables and institutional frames of a specific *context of situation* (Middle Range Theories, meta-level)
- (4) the broader socio-political and historical contexts, to which the discursive practices are embedded in and related (Grand Theories, meta-level).

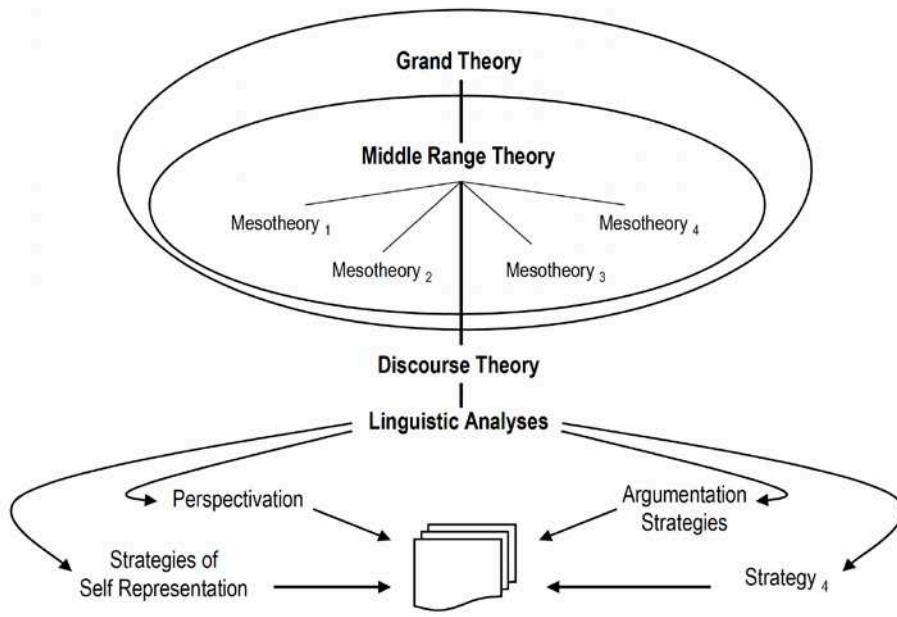


Figure 4. Meta-level theories of discourse: *Grand theory*, Middle-range and object-level Linguistic analyses.

Mesotheory delineates an intermediate level of analysis positioned between macro-theory, which scrutinizes large-scale social structures and institutions, and micro-theory, which delves into individual-level interactions and behaviors. Its focus is on understanding and scrutinizing social units of intermediate size, such as organizations, communities, or small groups.

In sociology and other social sciences, researchers employ meso-level analysis to investigate the relationships, patterns, and dynamics within these intermediate-sized units. This analytical approach facilitates a more nuanced comprehension of how social structures and individual behaviors intersect and mutually influence one another. Mesotheory acts as a bridge, connecting the broader social context with the specific interactions of individuals.

For instance, a sociologist may employ meso-theory to examine the dynamics within a particular workplace, exploring how organizational structures and group interactions impact the behavior and experiences of employees. Similarly, in community studies, mesotheory could be utilized to investigate how community organizations and local institutions contribute to shaping the social life of residents.

Fields of action, drawing from Pierre Bourdieu's concept of the field, can be conceived as segments of societal reality that play a role in constituting and shaping the framework of a discourse. In the domain of political action, distinctions can be made between various functions such as legislation, self-presentation, the formation of public opinion, internal party consensus-building, advertising and campaigning, governance and execution, as well as oversight and expression of oppositional dissent (Figure 5).

In the domain of sociolinguistics, Labov and Waletzky (1967) laid the groundwork for understanding the prevalent structure of oral narratives through their analysis of stories obtained during interviews with informants. This seminal work has left a lasting impact on both sociolinguistics and discourse studies. Essentially, Labov and Waletzky deconstructed each narrative into distinct components: orientation (the introduction), exposition (events introduction), culmination (the pivotal point of the story, often a surprising occurrence), and coda (the moral, summarizing the story). While subsequent studies have provided nuanced elaborations on this basic structure, the fundamental narrative line appears to align with both spontaneous and fictional storytelling, as demonstrated by Schiffrin (1994). This shift in focus extends from inherent textual characteristics to a more functional approach and eventually towards an examination of social

practices, conventions, and norms that govern specific sets or groups of speakers and listeners (viewers). For instance, a policy paper addressing unemployment is a manifestation of specific rules and expectations dictated by social conventions. The proposal adheres to particular textual devices, while its contents reflect ideological concepts advocated by a specific political group, such as trade unions. In the context of this chapter, discourse is expressed as a graph structure reflecting either thought process or interaction of parties of a healthcare encounter.

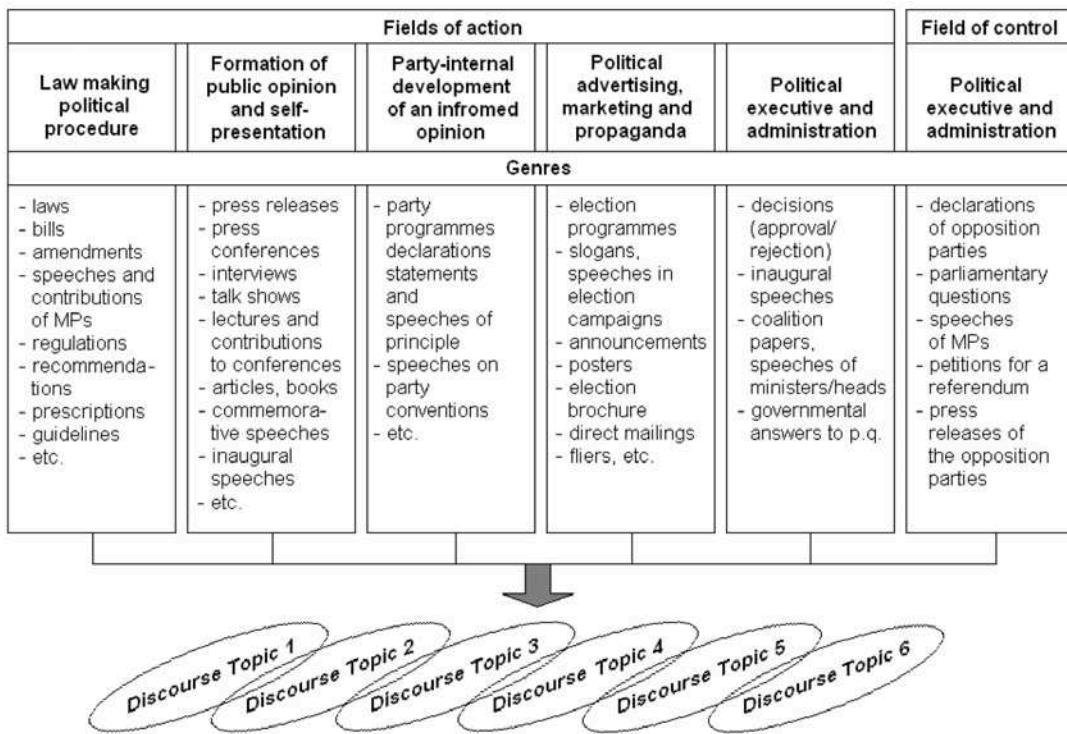
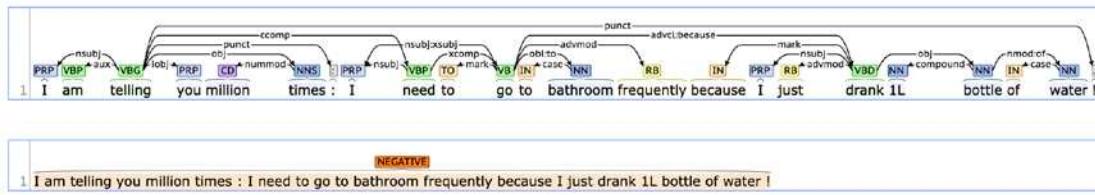


Figure 5. Selected dimensions of Discourse as Social Practice.

1.4. Why Discourse Analysis is Needed to Supplement LLM

We now show an example of question where a reasoning is required to answer it. We try to make a diagnosis, formulating a question “Why does a patient urinate too frequently?” We use an example patient’s complaint “I am telling you million times: I need to go to bathroom frequently because I just drank 1L bottle of water”. We build a linguistic and knowledge dossier below so that it become clear which representation layer is essential.

Our syntactic representation is as follows:



Notice the overall negative sentiment value associated with text.

Semantic parsing for “Why does a patient urinate too frequently?”

(mu / multi-sentence

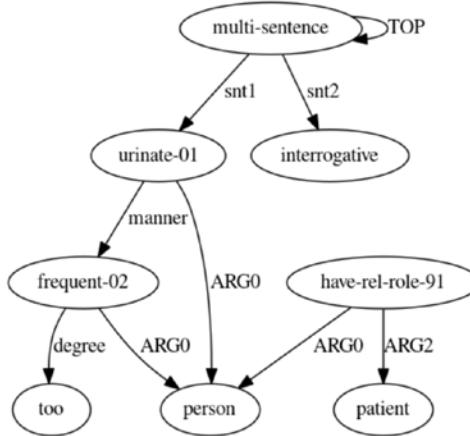
:snt1 (v4 / urinate-01

:ARG0 (v1 / person

:ARG0-of (v2 / have-rel-role-91

:ARG2 (v3 / patient)))

```
:manner (v6 / frequent-02
  :degree (v5 / too)
  :ARG0 v1))
:snt2 interrogative)
```



Background info:

Frequent urination, also known as urinary frequency, can be caused by various medical conditions. It's important to note that frequent urination can be a symptom of different underlying issues, and a healthcare professional should be consulted for an accurate diagnosis. Some common medical conditions associated with frequent urination include:

Urinary Tract Infection (UTI): Infections in the urinary tract, such as a bladder infection, can cause frequent urination along with other symptoms like pain or burning during urination.

Diabetes: High blood sugar levels associated with diabetes can lead to increased thirst and, consequently, frequent urination.

Enlarged Prostate: In men, an enlarged prostate gland can obstruct the flow of urine, causing more frequent urination.

Interstitial Cystitis: This chronic condition can cause bladder pain and an urgent need to urinate frequently.

Overactive Bladder (OAB): OAB is a condition characterized by a sudden, uncontrollable urge to urinate, leading to increased frequency.

Bladder or Prostate Cancer: Tumors in the bladder or prostate can cause changes in urinary habits, including increased frequency.

Anxiety or Stress: Emotional factors, such as anxiety or stress, can sometimes lead to increased urinary frequency.

Default ChatGPT answer which does not take into account generalized discourse: that the patient *is in distress*, according to how he writes:



You

"I am telling you million times!!!!: I need to go to bathroom frequently because I just drank 1L bottle of water!" what are the reasons for frequent urination?



ChatGPT

Frequent urination can be caused by various factors, and in the context you provided, the primary reason appears to be the consumption of a large volume of water.

However, persistent or unexplained frequent urination may be indicative of an underlying medical condition. Here are some common reasons for frequent urination:

1. **Excessive Fluid Intake:** Consuming a large amount of fluids, as in your example, can lead to increased urination.
2. **Urinary Tract Infection (UTI):** Infections in the urinary tract, including the bladder, can cause frequent and urgent urination.

Discourse representation:

elaboration

TEXT:I am telling you million times :

background

TEXT:I need to go to bathroom

explanation

TEXT:frequently

TEXT:because I just drank 1L bottle of water

We now explain how retrieval-augmented generation can leverage LLM. Discourse tree helps to find the relevant portion of text. Instead of submitting all paragraphs to LLM, the relevant paragraph is identified by a discourse tree (Figure 6)

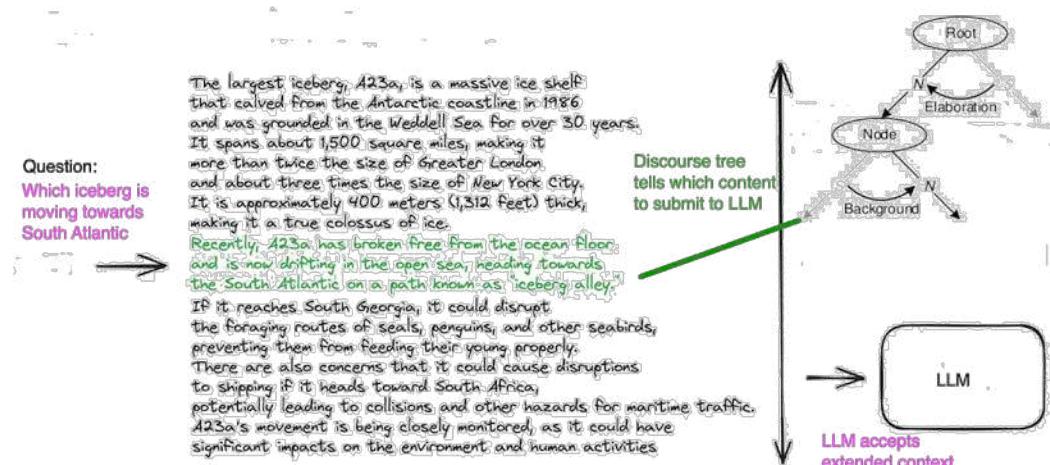


Figure 6. Sentence window retrieval.

The default approach is to retrieve smaller chunks for better search quality, but add up surrounding context for LLM to reason upon. This is done by either expanding context by sentences around the smaller retrieved chunk or by splitting documents recursively into a number of larger parent chunks, containing smaller child chunks. Each sentence in a document is embedded separately which provides great accuracy of the query to context cosine distance search. In order to better reason upon the found context after fetching the most relevant single sentence we extend the context window by k sentences before and after the retrieved sentence and then send this extended context to LLM.

Our proposal in this chapter is to rely on discourse tree, so that we know exactly which chunks are needed to answer a question.

1.5. Contribution

We enumerate the contribution of this chapter:

- (1) We developed the *MedDiscourse* system to answer queries for both unstructured and structured medical documents by harnessing LLMs.
- (2) In constructing MedDiscourse, we delved into the medical discourse literature, exploring the potential application of its features for responding to inquiries within lengthy and intricate medical documents, such as electronic health records. Our focus included a thorough examination of dialogue logs with patients, leading to the development of a discourse model tailored specifically for the medical domain.
- (3) Within our discourse model, we integrated the structure of patient interviews, adeptly handled metaphoric language used by patients, addressed various communication modalities found in text, and implemented a specialized discourse mechanism to represent pain.
- (4) Expanding beyond the conventional notion of a discourse tree, we broadened our model to encompass the entirety of a document, reflecting the diverse text structures found in genres ranging from diagnosis-making to treatment plans.
- (5) Effective discourse analysis requires an understanding of the social context in patient-doctor interactions to filter out response candidates influenced by social norms rather than valid medical information. Acknowledging the unique aspects of online doctor-patient communication, including motivations and trust considerations, we tailored our approach to provide pertinent answers and identify the root causes of issues.
- (6) Through our exploration, we discovered that discourse cues can reveal concealed or implicit data during the diagnostic process, compensating for missing information in the text. Overall, we observed that addressing discourse challenges can serve as a substitute for the absence of common sense and medical knowledge required to answer questions that demand a deep understanding of lengthy documents with varied structures.
- (7) Our proposed approach adopts a neuro-symbolic paradigm, where the LLM serves as the baseline for question-answering, and discourse analysis operates at the symbolic level, effectively "spreading" question-answering capabilities across lengthy, unstructured documents.

The primary aim of this book is to captivate readers by exploring diverse applications of discourse analysis. Each chapter will delve into a specific application, and in this particular chapter, we delve into the establishment of connections between documents pertaining to individual diseases, forming a multi-case for Case-Based Reasoning (CBR). Despite significant progress in document-level Rhetorical Structure Theory (RST) parsing, incorporating feature-rich linguistic parsing models (Joty et al. 2013), the broader application of document-level discourse analysis remains limited. Extracting valuable insights from Discourse Trees (DT) involves considering global discourse features and the long-range structural dependencies between DT constituents.

The study of discourse trees and their extensions holds considerable promise within the realm of logical Artificial Intelligence (AI). Logical AI focuses on subjects like logic forms and logic programs, which are relatively scarce in the real world. However, discourse trees offer fairly interpretable structures, making them suitable for logical AI exploration. While statistical/deep

machine learning has access to vast amounts of text data, it often falls short of making sense of this data from the perspective of Logical AI. Communicative discourse trees present an opportunity for abundant acquisition and serve as suitable subjects for Logical AI, making them a crucial area of study, especially in domains like healthcare.

2. Extending the Model of Based on Phenomenology of Medical Discourse

2.1. Discourse Analysis and Discourse Trees

Discourse analysis of text, particularly through communicative discourse trees, aims to integrate rhetorical information with speech act structures. A Discourse Tree (DT) is a hierarchical representation that captures the organizational and structural relationships among elements in a discourse or text. It is a graphical structure where nodes represent elementary discourse units (EDUs) or larger discourse segments, and edges between nodes indicate the rhetorical or discourse relations between them.

In the context of discourse analysis, a DT helps visualize how different parts of a text are interconnected and how they contribute to the overall meaning and coherence. The tree structure allows for the representation of rhetorical relations, indicating how one segment of text relates to another in terms of functions like elaboration, contrast, cause-effect, and more. Discourse Trees are commonly used in computational linguistics, natural language processing, and discourse analysis to study the organization of information and the flow of ideas within a text. They provide a visual tool for understanding the discourse structure and can be instrumental in tasks such as text summarization, information retrieval, and sentiment analysis.

Communicative DTs (CDTs), are DTs with labeled arcs denoting expressions for communicative actions, utilize logic predicates to represent agents involved in speech acts and their subjects. These logical predicates follow semantic roles proposed by frameworks like VerbNet (Kipper et al., 2008), enriching DTs with speech act-specific details beyond rhetorical relations and the syntax of elementary discourse units (EDUs). This approach comprehensively captures how authors organize and convey thoughts, irrespective of the subjects involved (Galitsky, 2017).

The key discourse connections between sentences include:

- (1) Anaphora: Keyword occurrences in two areas connected by an anaphoric relation suggest relevance, enhancing the likelihood of a pertinent answer.
- (2) Communicative Actions: In a dialogue, if question keywords are present in a doctor's question and others in the patient's reply, connecting these keywords establishes relevance. Identifying such situations involves confirming that a pair of communicative actions is of the question-answer or request-reply type (Galitsky and Kuznetsov, 2008; Galitsky, 2019a).
- (3) Rhetorical Relations: These relations signify the coherence structure of a text (Mann and Thompson, 1988). Represented by a DT, rhetorical relations organize adjacent EDUs and higher-level discourse units in a hierarchy based on relation types (e.g., Background, Attribution). Anti-symmetric relations involve pairs of EDUs, including nuclei (core parts) and satellites (supportive parts).

In this book, we primarily focus on the crucial discourse connection class between sentences, namely rhetorical relations. After splitting an answer text into Elementary Discourse Units (EDUs) and establishing rhetorical relations between them, we can formulate rules to determine whether query keywords in the text are connected by rhetorical relations. This process helps identify relevant answers (connected) or irrelevant answers (not connected). By employing Discourse Trees (DTs), specific sets of nodes correspond to valid answers, while others correspond to invalid ones.

Discourse parsing, essential for obtaining DTs from text, is a complex challenge that requires the understanding and modeling of various semantic and pragmatic features. Additionally, it involves grasping the structural properties inherent in a DT. Many current theories and computational models present a simplified version of discourse structure. For instance, Rhetorical Structure Theory (RST, Mann and Thompson, 1988; Taboada and Mann, 2006) stipulates that only adjacent EDUs should be connected with a rhetorical relation, illustrating how a text author organizes their thoughts.

Another prevalent discourse model, the Penn Discourse Treebank (PDTB, Prasad et al., 2008), addresses the discourse connectives' attachment issue but doesn't impose constraints on the overall discourse structure in the resulting annotation. Computational models of PDTB simplify the attachment problem, making it suitable for a broad range of Natural Language Processing (NLP) tasks.

2.2. Forming a Discourse Tree for a Health Complaint

We take a typical patient complaint and build a discourse tree for it (Figure 6, PatientInfo 2023).

I am 3 months out from Total Heap Replacement and my leg is still hard to lift and get into car or lift onto bed or couch. I have to help lift it. I understand the feeling of lifting the operated leg: it can be tricky during the first few months of recovery. This is because the surgery is a trauma on the body. Therefore, it takes time for the entire leg to regenerate and heal. So many structures have been effected along with the mechanics of movement, muscles, soft tissue and bones to heal . Surgeons remove the entire joint then replaced with a prothesis: it is a major surgery. It is painful. I am taking pain control drugs. My doctor is saying I can still perform my housewife responsibilities by doing stuff for which I do not have to walk. I contacted the hospital on the internet to get some support, but when I described my problems I was always interrupted. I am early in the short-term healing process: it is usually 3-6 months. Notice that long term healing is 6 month to a year. I have to hang in there; it will get better as my body continues to heal and gets stronger. Soon you will have your full mobility back and a better quality of life.

elaboration

explanation

elaboration

elaboration

joint

TEXT:I am 3 months out from Total Heap Replacement [stage-medical-encounter]

TEXT:and my leg is still hard to lift and get into car or lift onto bed or couch .

TEXT:I have to help lift it .

elaboration

TEXT:I understand the feeling of lifting the operated leg :

TEXT:it can be tricky during the first few months of recovery .

elaboration

TEXT:This is because the surgery is a trauma on the body .

elaboration

TEXT:Therefore , it takes time for the entire leg

TEXT:to regenerate and heal .

elaboration

TEXT:So many structures have been effected along with the mechanics of movement , muscles , soft tissue and bones to heal .

elaboration

elaboration

elaboration

elaboration

elaboration

TEXT:Surgeons remove the entire joint [stage-medical-encounter]

TEXT:then replaced with a prothesis :

TEXT:it is a major surgery .

TEXT:It is painful . [pain discourse]

TEXT:I am taking pain control drugs . [pain discourse]

elaboration

attribution (RightToLeft)

TEXT:My doctor is saying

manner-means

TEXT:I can still perform my housewife responsibilities [social]

elaboration

TEXT:by doing stuff

TEXT:for which I do not have to walk .

elaboration

contrast (RightToLeft)

elaboration

TEXT:I contacted the hospital on the internet [online communication]

TEXT:to get some support ,

same-unit

TEXT:but

background (RightToLeft)

TEXT:when I *described* my problems [**communicative action**]

TEXT:I was always interrupted [patient-doctor dialogue structure]

elaboration

elaboration

TEXT:I am early in the short-term healing process :

TEXT:it is usually 3-6 months .

elaboration

TEXT:Notice that long term healing is 6 month to a year .

elaboration

elaboration

TEXT:I have to hang in there ;

background

TEXT:it will get better

joint

TEXT:as my body continues to heal

TEXT:and gets stronger .

TEXT:Soon you will have your full mobility back and a better quality of life .

Figure 6. Discourse tree with health-specific labels.

In Figure 6, we tag [stage-medical-encounter], [medical encounter], [patient-doctor dialogue structure], [social], [online communication], and [pain discourse] labels. Communicative actions are shown in *italic*[communicative action]. All this information is essential to do chunking and do a special discourse-oriented index. Complete representation of medical discourse turns out to be essential for asking complex health-related questions which require reasoning.

2.3. Additional Health-Specific Labels in Discourse Representation

On top of logical organization of a paragraph or a document containing medical text of various genre, following the corpus of literature on medical discourse, we add the health-specific labels:

- (1) Stage in the medical encounter process (Chief complaint (CC), present illness(PI), past history (PH), family history (FH), social history (SH), systems review (SR), physical examination (PE), other investigations, diagnosis(Dx), plan (P), and recovery in Sect... [stage-medical-encounter].
- (2) Discourse markers of questioning, interrupting, shifting the direction of conversation and other dialogue-based modifications in Sect ..., [patient-doctor dialogue structure],
- (3) Ideology and social control markers in Sect ... [social],
- (4) Pain management discourse (Sect ...) is marked with [pain discourse],
- (5) Online communication components (Sect) [online communication].
- (6) Handling nontechnical, nonmedical problems that patients bring into the medical encounter.

All these labels are need to perform some commonsense reasoning steps.

3. Answering Questions Based on Document Discourse

We now turn our attention to the challenge of evidence retrieval for answering questions in the context of long medical documents. This task involves identifying and selecting relevant paragraphs within a document that contain information necessary to address a given question. The difficulty arises from the fact that lengthy documents often surpass the token limit of current transformer-based Pretrained Language Models. Directly processing the content of these documents to extract pertinent information becomes a challenge. Additionally, the required information for answering a question is often distributed across various sections or paragraphs, necessitating advanced reasoning processes for identification and extraction of the pertinent details (Nie et al., 2023). Attempting to process the entire document to find answers without leveraging its discourse structure can be both computationally expensive and inefficient.

3.1. Employing Document Structure

The cognitive strategy utilized by humans to locate pertinent information in a document involves a systematic approach. Initially, individuals categorize the information within the document to identify relevant coarse segments. Subsequently, they delve deeper into the relevant categories to conduct a more detailed analysis and extract fine-grained segments.

Instead of representing a document D as an ordered set of constituent paragraphs, we represent $D = [S_1, S_2, \dots, S_k]$, where S_i ($1 \leq i \leq k$) denotes section, such that, $name(S_i)$ and $paragraphs(S_i)$ denotes its name / heading and the list of constituent i^{th} paragraphs respectively.

$paragraphs(S_i) = [p_i, j]_{j=1}^{|S_i|}$, where $|S_i|$ denotes number of constituent paragraphs). Note that $\sum_{i=1}^k |S_i| = n$. Following the cognitive process of knowledge acquisition / information search for question answering, the proposed approach first finds the relevant sections that may answer the question and then, analyses the paragraphs from the relevant sections for fine-grained evidence paragraph retrieval (Nair et al 2023).

Documents often exhibit a hierarchical discourse structure, encompassing various levels of sections (Nair et al., 2023). To address this, the structure can be flattened by employing a preorder traversal approach. When expressing a particular section, we concatenate the names of all sections along the path from the root node to that specific node in the discourse structure. This flattening technique enables us to represent the document as a list of sections, considering the hierarchical

relationships among them. Flattening a tree using a preorder traversal involves visiting each node in a specific order—starting with the root node, followed by recursive traversal of the left subtree, and concluding with recursive traversal of the right subtree. The entire document is represented as

Upon receiving a question and a lengthy document featuring the results of extended discourse parsing that denotes sections, subsections, etc., the task involves pinpointing the pertinent sections necessary to address the question (Figure 7). Following this, relevant paragraphs are extracted from the narrowed-down list of paragraphs within the relevant sections. In the third step, these identified paragraphs are then fed into a Language Model for question answering.

- * Section: $\text{name}(S_1)$
 $\mathcal{S}(\text{paragraphs}(S_1))$
- * Section: $\text{name}(S_2)$
 $\mathcal{S}(\text{paragraphs}(S_2))$
- ...
- * Section: $\text{name}(S_k)$
 $\mathcal{S}(\text{paragraphs}(S_k))$

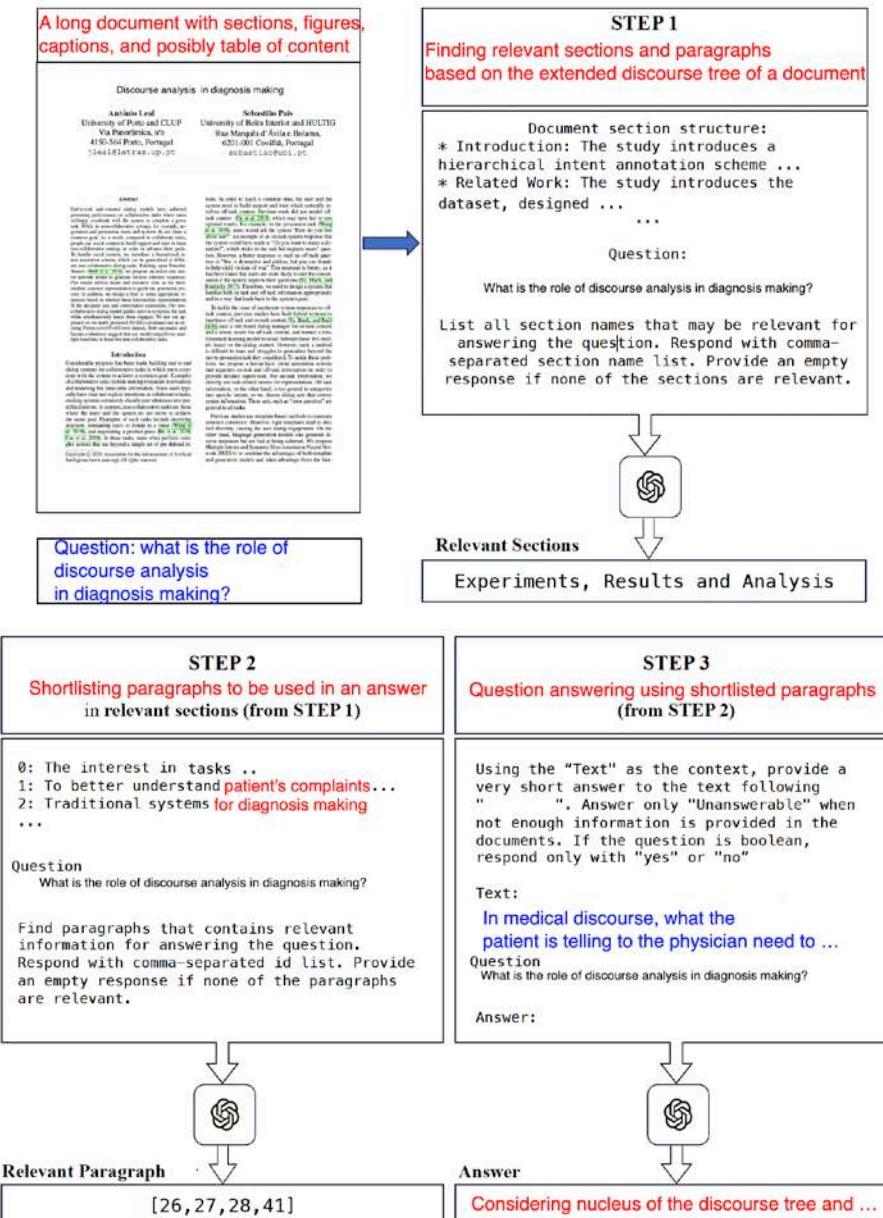


Figure 7. Steps for finding an answer in a long document with varied structure.

The annotated question-summary hierarchy for sentences in a reference summary paragraph is illustrated in Figure 8. Summarization models undergo training to produce the question-summary hierarchy based on the document, emphasizing the significance of encoding the document structure. For example, generating follow-up question-summary pairs like Q1.1 and A1.1 from A1 necessitates understanding both the content and the relationships among §3, §3.1, and §3.4, including parent-child and sibling relations.

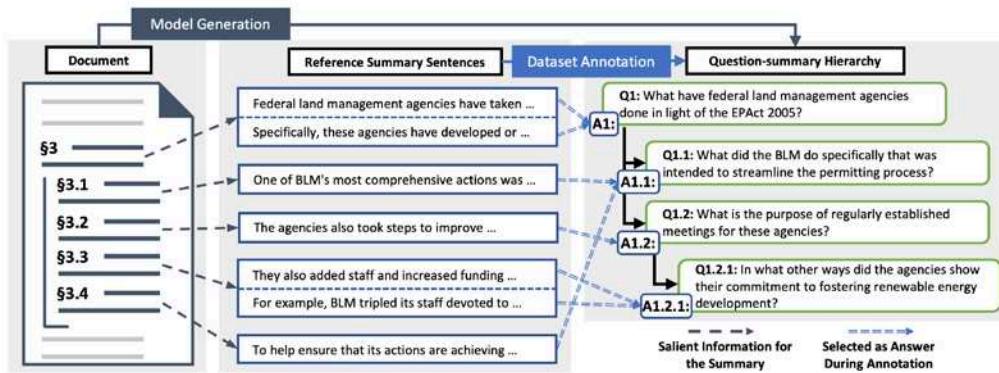


Figure 8. Handling a hierarchical organization of document.

Leveraging section structures, Cohan et al. (2018) devise a section-level encoder utilizing the output of a word-level encoder for long document summarization. However, multi-level encoders incur higher costs as they introduce a substantial number of parameters and additional padding at various model levels. In contrast, Cao and Wang (2022) effectively incorporate document structure information by introducing a novel bias term in attention calculation among tokens, which introduces only a small number of learnable parameters.

In the field of Long Document Summarization, the inclusion of document structure information proves advantageous. Extractive summarization methods aim to amalgamate section-level and sentence-level information encoded by multilevel encoders (Xiao and Carenini, 2019) and incorporate longer context through sliding encoding over sections (Cui and Hu, 2021). Recent advancements in summarizing long documents focus on designing efficient Transformers with sparse attentions to generate abstractive summaries in an end-to-end manner (Beltagy et al., 2020; Zaheer et al., 2020; Huang et al., 2021). However, these approaches often overlook the natural structure of long documents, such as sections and subsections. Cao and Wang (2022)'s system, based on a simple design, seamlessly integrates into any efficient Transformer, facilitating the incorporation of document structure information.

3.2. Discourse-Free Approach to Long Document QA

Nie et al (2023) propose a new task, named unsupervised long-document question answering, aiming to generate high quality long-document QA instances in an unsupervised manner. Besides, we propose a novel unsupervised attention-walking method to aggregate and generate answers with long-range dependency so as to construct long-document QA pairs. Proposed system is composed of three modules (Figure 9), *EDU collector*, *EDU linker* and *Answer fusion*.

- (1) The EDU collector takes advantage of constituent parsing and reconstruction loss to select informative candidate spans for constructing answers.
- (2) By going through the attention graph of a pre-trained long-document model, potentially interrelated EDUs (that might be far apart) could be linked together via an attention-walking algorithm.
- (3) In the Answer fusion component, linked EDUs are aggregated into the final answer via the mask-filling ability of a pre-trained model.

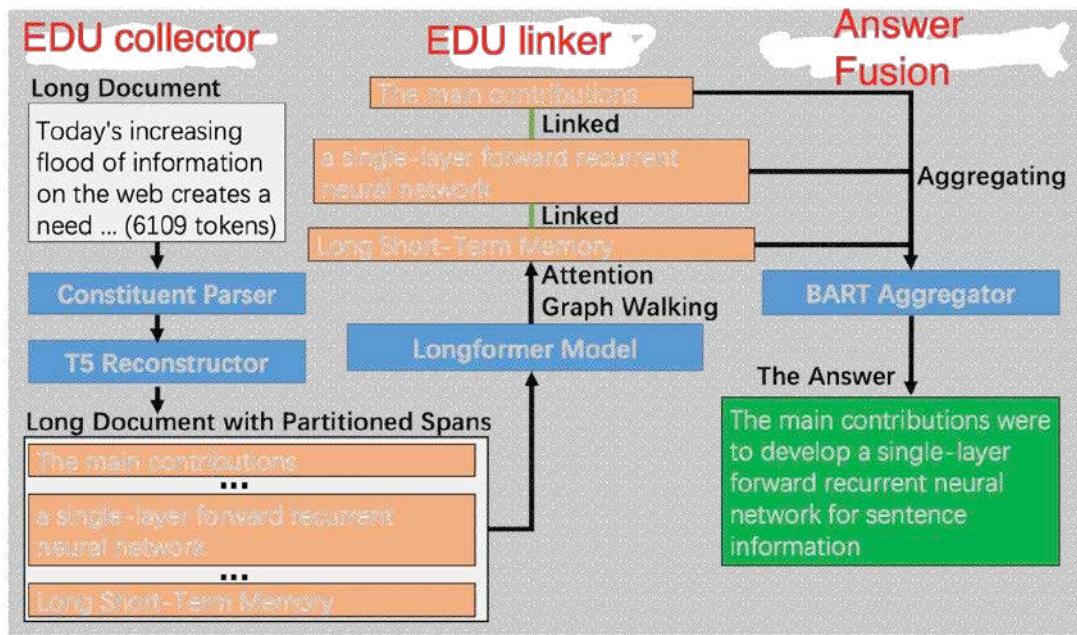
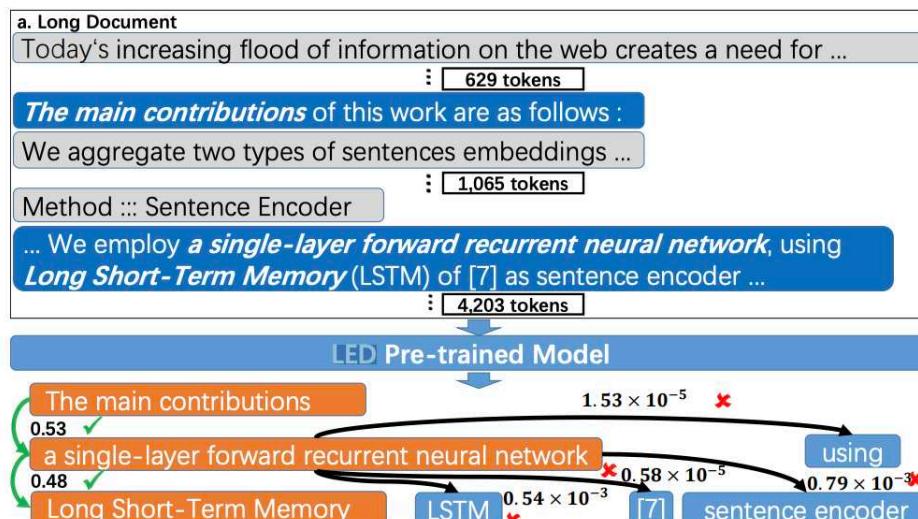


Figure 9. An architecture for unsupervised long-document question answering.

The process of discovering long-range relations in a long document is illustrated in Figure 10. Initially, the document undergoes processing by a long-document encoder-decoder pre-trained model (Beltagy et al., 2020) (depicted in the upper part of the figure). Subsequently, the token-level attention graph, although not explicitly shown here, is transformed into a span-level graph (depicted in the lower half). Spans, which may be widely separated, are connected if the weight of their edge is high. For instance, the span "The main contributions" traverses a thousand tokens and forms links with "a single-layer forward recurrent neural network," which, in turn, connects with "Long Short-Term Memory" due to their high-weight edges (0.53 and 0.48 in this example). Other spans do not establish connections with them because of the low edge weights associated with these spans.

To enhance attention optimization for longer documents, the following strategies are employed to circumvent the computation of the entire attention matrix:

- (1) Tokens attend to each others following an "attention pattern";
- (2) Large receptive field with stacked layers.



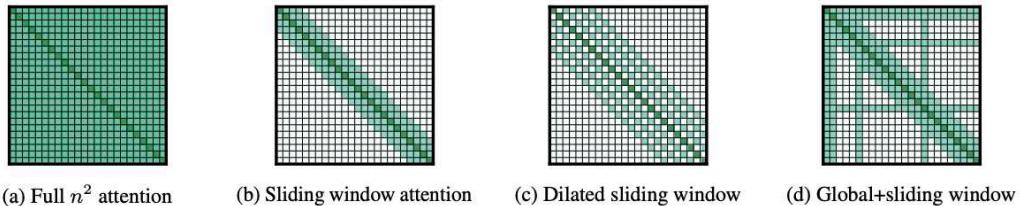


Figure 10. A long-range relation discovering process (on the top). Efficient attention patterns for Transformers (on the bottom, (Beltagy et al., 2020)).

3.3. Prompt-Based Approaches

(Saad-Falcon et al., 2023) employs the structured metadata of a PDF to enhance the precision and accuracy of a document question-answering system. The process begins with the generation of a structured metadata representation, capturing details from section text, figure captions, headers, and tables. When presented with a query, a LLM-based system identifies the relevant document frame for answering the query and directly retrieves it from the chosen page, section, figure, or table. Subsequently, the LLM processes the inputted query and the selected context, producing the final answer.

The method involves three key steps in addressing user questions (refer to Figure 11):

- (1) Generate Document Metadata: Extract structural elements from the document and convert them into readable metadata. Utilizing the Adobe Extract API, a PDF is transformed into an HTML-like tree, facilitating the extraction of sections, section titles, page details, tables, and figures. The tree is parsed to identify sections, section levels, headings, and gather text from specific pages, figures, and tables. This structured information is then mapped into a JSON format, serving as the initial input for the LLM.
- (2) LLM-based Triage: Query the LLM to pinpoint precise content (pages, sections, retrieved content) from the document, focusing on structured textual data in headers, sub-headers, figures, tables, and section paragraphs. Individual queries are formulated for each question, integrating multiple pieces of information to derive the ultimate answer. Answer using retrieved content: Based on the question and retrieved content, generate an answer. The following prompt is used: "You are an expert document question answering system. You answer questions by finding relevant content in the document and answering questions based on that content. Document: {textual metadata of document}"

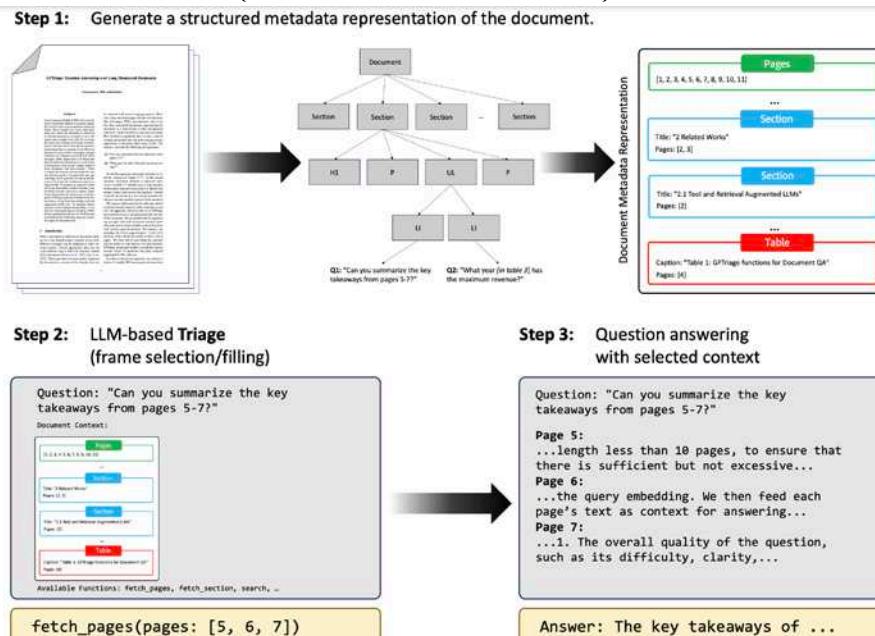


Figure 11. Prompt-based answering questions about long documents.

3.4. Embedding Discourse Tree

(Cao and Wang, 2022) introduces learned biases in attention weight calculation to integrate hierarchical document structure, leading to improved summarization of long documents. This underscores the significance of considering hierarchical document structure for a comprehensive understanding of lengthy documents. Conversely, (Du et al., 2023) emphasizes section-level structural relations, such as parent-child and sibling connections, focusing on aspects like token-level path lengths and level differences within the document structure graph.

The architecture proposed by (Du et al., 2023) comprises four main components (Figure 12):

- (1) Contextual Encoder;
- (2) Sentence-level Discourse Graph Encoder;
- (3) Section-level Structure Graph Encoder;
- (4) Fusion and Decoding.

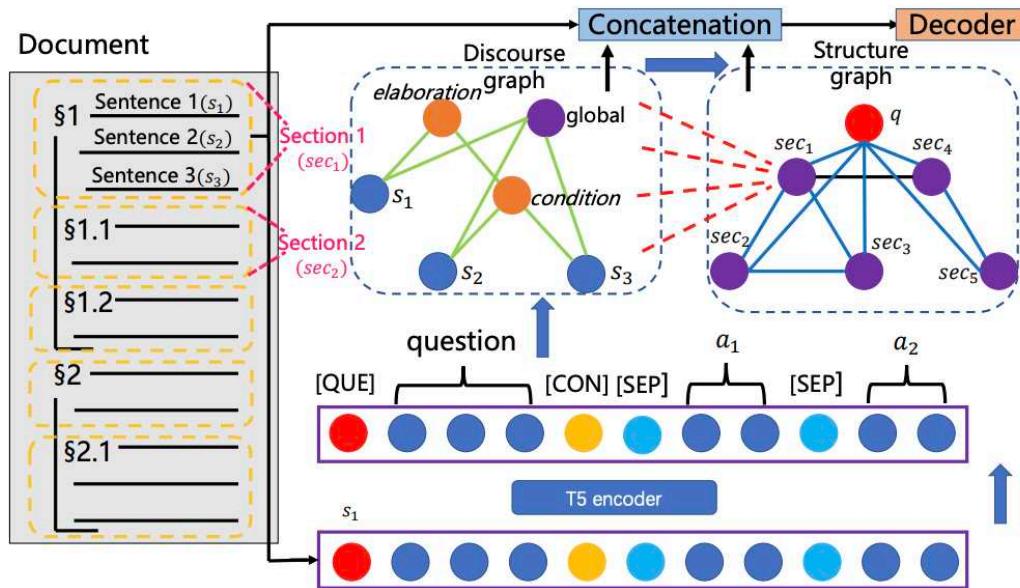


Figure 12. Encoding discourse tree into the document-level.

The hierarchical graph facilitates information propagation from the bottom to the top, where "elaboration" and "condition" represent two types of rhetorical relations (Figure 13).

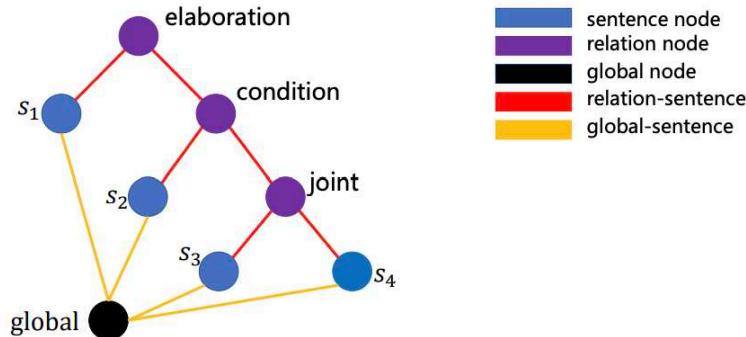


Figure 13. hierarchical discourse super-graph.

4. Sentence- and Section-Level Discourse Graph Encoder

The discourse tree consists of two main node types: relation nodes and leaf nodes. In this context, a leaf node represents a sentence within the section, while a relation node identifies the relation type

between two consecutive text spans. To convert the discourse tree into a discourse graph for each document section, we introduce a global node connected to every sentence node.

The generative model processes the concatenation of the question and context as input to derive contextual representations. Specifically, each document section, concatenated with the question, is independently processed by the encoder. Special tokens "[QUE]" and "[CON]" are added before the question and context, respectively, along with "[SEP]" to separate each sentence in the document section. This approach is scalable to long documents with multiple sections as it encodes one section at a time.

For each document section, a discourse relation graph is constructed to incorporate relational information between sentences within the section (Du et al 2023). The RST discourse parser is utilized to derive the discourse graph, including a global node representing the entire section. Edges are added between the global node and the leaf nodes (sentence nodes) to enhance information flow among sentences. For a section comprising n sentences, the representation of the global node is initialized as the hidden state h_0 corresponding to "[CON]" from the contextual encoder, and the representation of the t -th sentence node as the hidden state corresponding to the t -th "[SEP]" token h_t , $1 \leq t \leq n$. The Graph Attention Network is employed for information propagation and to derive discourse relation-enhanced representations.

$$\begin{aligned} \mathbf{a}_{ij} &= \text{MLP}([\mathbf{h}_i : \mathbf{h}_j]) \\ \alpha_{ij} &= \frac{\exp(\text{LeakyRelu}(\mathbf{a}_{ij}))}{\sum_{j' \in N(i)} \exp(\text{LeakyRelu}(\mathbf{a}_{ij'}))} \\ \hat{\mathbf{h}}_i &= \sigma \left(\sum_{j \in N(i)} \alpha_{ij} \mathbf{W} \mathbf{h}_j \right) \end{aligned}$$

where $N(i)$ denotes the neighbor nodes of node i , $1 \leq i \leq n$ and σ denotes activating function. We take the final representation of global node as the section representation that incorporates the discourse relational information.

4.1. Section-Level Document Graph Encoder

We create a node for each document section in the structure graph and introduce the question node. Section nodes are linked to their parent section and child sections, capturing the information that a child section pertains to a specific aspect of the parent section. Nodes at the same level with the same parent section are connected as siblings, signifying that these sections elaborate on parallel and relevant aspects of the parent. Additionally, we establish connections between the question node and each section node to facilitate information flow between the question and contexts.

The question representation, q , is initialized as the hidden state corresponding to the "[QUE]" token obtained in the contextual encoder module. The initialization of section nodes is derived from the representations of global nodes in the corresponding DTs. Information transmission on the structure graph is facilitated by the Graph Attention Network, resulting in structure-aware section representations:

$[q'; h'_1; h'_2; \dots; h'_N] = \text{GraphAttentionNetwork}([q; h_1; h_2; \dots; h_N])$, for N sections in the document.

Graph Attention Networks use attention mechanisms to assign different importance scores to neighboring nodes during the aggregation of information. A simplified pseudocode is shown in Fig ... x represents the node features, and edge_index represents the adjacency matrix.

The DT needs to be pre-processed to create the necessary inputs for the GAT model. To handle variations in the discourse tree structures, since Graph Attention Networks assume a fixed graph structure, the Graph Attention Networks needs to be padded or truncated to a consistent size (Figure 14)

```
class DiscourseGAT(nn.Module):
    def __init__(self, num_nodes, node_features, num_classes):
        super(DiscourseGAT, self).__init__()
        self.gat_conv1 = GATConv(node_features, 64, heads=8, dropout=0.6)
```

```

self.gat_conv2 = GATConv(64 * 8, num_classes, heads=1, concat=True, dropout=0.6)

def forward(self, x, edge_index):
    x = F.dropout(x, p=0.6, training=self.training)
    x = F.elu(self.gat_conv1(x, edge_index))
    x = F.dropout(x, p=0.6, training=self.training)
    x = self.gat_conv2(x, edge_index)
    return F.log_softmax(x, dim=1)

# Instantiate the model
num_nodes = 10  # Adjust this based on the number of nodes in your discourse tree
node_features = 64  # Adjust this based on the dimensionality of your node features
num_classes = 2  # Adjust this based on the number of classes in your classification task
model = DiscourseGAT(num_nodes, node_features, num_classes)

```

Figure 14. Pseudo-code for discourse Graph Attention Network.

4.2. Graph Decoding

Hence we have contextual information $\{h_i, 1 \leq i \leq n\}$, discourse relational information $\{h_i, 1 \leq i \leq n\}$, including health-related labels (Section 2.3) and document structure information, $\{h'_i, 1 \leq i \leq n\}$. The token representations of question and contexts are concatenated, as well as three levels of section representations sequentially as follows:

$$[t_1^q; t_2^q; \dots; t_Q^q; t_1^c; t_2^c; \dots; t_C^c; h_1; h_2; \dots; h_N; \bar{h}_1; \bar{h}_2; \dots; \bar{h}_N; \bar{h}'_1; \bar{h}'_2; \dots; \bar{h}'_N]$$

where Q, C, N denote the number of question tokens, context tokens and document sections correspondingly. Then we pass them into PLM decoder to generate the sequence shaped as

“... [ANS] a_i [CON] $c^{(i)}$... [CON] $c^{(i)}_{N(i)}$...” where a_i and $c^{(i)}$ denote the i -th answer and the j -th conditional rhetorical relation of the i -th answer, “[ANS]” and “[CON]” are special tokens added into LLM tokenizer. The model is optimized by the cross-entropy loss between the predicted sequence and ground truth

$$\begin{aligned}
L &= -\log p(r|q, C) \\
&= -\sum_{i=1}^L \log p(r_i|q, C, r_{<i})
\end{aligned}$$

where $r = \{(a_i, c_i)\}_{i=1}^L$, a_i and c_i denote the i -th answer and conditional rhetorical relation, used as an example of rhetorical relations.

5. Evaluation

5.1. Datasets

In this analysis, we utilize the HotpotQA-Doc dataset (Yang et al., 2018), which aims to address intricate queries involving multi-hop reasoning based on two lengthy documents. (Nair et al., 2023) explore the effectiveness of the zero-shot direct processing approach. The authors also experiment with self-ask-based processing (not compared with our system). Utilizing elicitive prompting power (Press et al., 2022), the authors implement the self-ask technique, breaking down a complex query into a series of simpler questions that collectively contribute to the final answer. Through iterative questioning, the agent analyzes prior answers and previously posed questions to generate

subsequent inquiries. Leveraging the zero-shot retrieval approach, the system derives relevant answers for each question.

The ConditionalQA dataset serves as a challenging benchmark for conditional QA over extensive documents (Sun et al., 2022), containing 3427 questions. The average document length exceeds 2K, as illustrated in Table 1. Table 2 showcases various question types, including yes/no questions, freeform extractive questions, questions with multiple answers, and not-answerable questions. Many questions in ConditionalQA are deterministic, where the necessary conditions are satisfied in the question.

The Qasper dataset (Dasigi et al., 2021) comprises information-seeking questions tailored for lengthy research papers. This dataset includes a set of ground truth evidence paragraphs and answers, with questions categorized as extractive, abstractive, yes/no, and unanswerable.

5.2. Answer Relevance

We compare our approach with five competitive approaches on long document QA:

- (1) ETC (Ainslie et al., 2020) applies global-local attention mechanism between global and local tokens, and enables the model scale to long inputs. However, the fully connected topology of token graphs cannot capture the natural structure of the document.
- (2) DocHopper (Sun et al., 2021) highlights the structural information that a passage contains consecutive and relevant information, and retrieves information by jointly sentence and passage level. However, the natural structural information between passages is ignored,
- (3) FID (Izacard and Grave, 2021) independently encodes different passages and concatenates the representations in the decoder only, which decreases calculation cost and improves performance for QA on long documents. However, the natural structure of documents and discourse information in each section are neglected.
- (4) SDHG (Structure-Discourse Hierarchical Graph, Du et al 2023) conducts bottom-up information propagation, firstly we build the sentence-level discourse graphs for each section and encode the discourse relations by graph attention. Secondly, a section-level structure graph is built based on natural structures, and conduct interactions over the question and contexts. Finally, different levels of representations are integrated into jointly answer and condition decoding.
- (5) D3 (Nair et al., 2023).

Table 1. Comparative performance of QA against long documents.

Dataset	HotpotQA-Doc		Qasper		ConditionalQ		
	Settings	Evidence	Answer	Extractive	Abstractive	Extractive	Conditional
gpt-3.5-turbo		41.0	54.9	27.8			
ETC					17.3	41.8	
DocHopper					26.7	46.4	
FID					37.8	49.7	
SDHD					42.0	52.3	
D ³	26.9	43.5	42.9	23.7			
MedDiscourse (ours)	23.2	42.0	56.4	24.7	44.2	47.1	

We show F1 accuracies answering questions. One can observe that the proposed system outperforms the other long-document QA in *ConditionalQ-Extractive* and *Qasper-Extractive* evaluation settings. For HotpotQA-Doc, the performance of D³ is systematically better. At the same time, SDHD shows a superior performance in the case of *Conditional* evaluation.

5.3. Answer Quality Scoring

We turn our attention to an evaluation conducted by annotators. In our annotation study, we tasked the annotators with ranking MedDiscourse in comparison to three baselines: Page Retrieval, Chunk Retrieval, and PDFTriage. The findings from (Saad-Falcon et al., 2023) indicate that annotators favored the PDFTriage answer over half of the time, and they also showed a preference for the Chunk Retrieval approach over the Page Retrieval approach. When comparing different provided answers for the same question, MedDiscourse demonstrated comparable performance to PDFTriage and significantly outperformed current alternatives, consistently ranking higher than the alternate approaches across all question types.

MedDiscourse not only enhances answer quality, accuracy, readability, and informativeness, but in our annotation study, annotators also assigned scores to PDFTriage, Page Retrieval, and Chunk Retrieval answers based on five major qualities: accuracy, informativeness, readability/understandability, and clarity. Table 2 illustrates that the MedDiscourse system, along with PDFTriage answers, ensures higher scores than Page Retrieval and Chunk Retrieval across all answer qualities except for Clarity.

Table 2. Improving the quality of answers relying on discourse.

	Readability	Informativeness	Clarity	Accuracy
Page Retrieval	4.1	3.7	2.1	3.6
Chunk Retrieval	4.1	3.4	2.3	3.4
PDFTriage	4.2	3.9	2.0	3.8
MedDiscourse	4.2	4.1	1.9	3.6

In Figure 15, the annotator is ready to receive a long question.



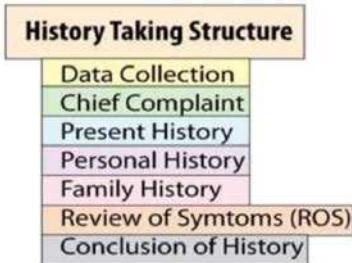
Figure 15. One of the annotators ready for testing.

6. Medical Encounter

We first start with analyzing discourse of a medical encounter in a general sense. Then we proceed to a patient-doctor dialogue (Sect. 7). We need this discourse analysis to answer questions concerning the diagnosis,

The traditional format of the medical encounter is as follows:

Chief complaint (CC) + present illness
(PI) + past history (PH) + family history
(FH) + social history (SH) + systems
review (SR) + physical examination (PE)
+ other investigations + diagnosis
(Dx) + plan (P).



In a standard encounter, the physician endeavors to address these components through verbal communication and the examination of the patient. Additionally, the physician documents the encounter in writing within the medical record, assigning each component with the provided abbreviations. Previously reported research on physician-patient communication has affirmed that medical practitioners do utilize the traditional structure as an organizational framework for their interactions with patients (Waitzkin 1985).

The components of the medical history (Hx) encompass CC, PI, PH, FH, SH, and SR in the aforementioned scheme. In the chief complaint (CC), the physician identifies the primary issue bothering the patient, succinctly and directly. Typically, the physician introduces the chief complaint with an opening question like, "Hello, where does it hurt?" or "Well, what is not giving you a good sleep?". In response, the patient might say, "migraine," "my knee hurts," "I've got pain in my stomach," "I can't sleep well," or "I want a check-up," among others. By eliciting the CC, the physician aims to understand the patient's primary concern.

In the course of the present illness (PI), the patient provides additional details regarding the chief complaint. This includes information on when the issue began, specific characteristics of the symptom, any medications or measures that alleviate the symptoms, previous medical attention received for the problem, and other details contributing to the physician's diagnostic efforts. Guiding the patient to articulate both the chief complaint and the present illness is regarded as a paramount skill that physicians develop in obtaining a medical history; some commentators even assert that it is the most crucial skill in medicine.

During the present illness discussion, interruptions by physicians often commence. Such interruptions are attempts to curtail the patient's narrative, driven by various reasons, including:

- (1) The patient's story may not contribute significantly to the physician's cognitive process of reaching a diagnosis.
- (2) The patient's version of the story may be confusing or inconsistent.
- (3) Narrating the story may exceed the perceived available time.
- (4) Parts of the story may evoke uncomfortable feelings for the physician, the patient, or both.

The circumstances surrounding physician interruptions during the patient's narrative in the present illness, such as what is interrupted, when it occurs, the reason given for the interruption, etc., hold significance, particularly in terms of potentially truncating discussions about the social context of the medical encounter.

Certainly, the present illness (PI) marks a pivotal moment wherein certain elements, despite their potential significance in the patient's experience, may be excluded from the discourse, while others are incorporated.

Although the chief complaint (CC) and present illness (PI) are virtually constant features in medical encounters, the inclusion or omission of other components depends on various factors such as time constraints, the physician's inclination to conduct a comprehensive evaluation, financial considerations like the patient's insurance coverage and the extent of evaluation permitted, and other situational limitations. A physician may opt to defer some or all of the remaining components for future visits or may choose not to address them at all, even though there is typically an initial effort to formulate a diagnosis and plan. During the past history (PH), the physician collects information about prior medical events in the patient's life that are not directly relevant to the present illness. These events commonly encompass previous hospitalizations and surgeries, significant illnesses, details about medications, allergies, immunizations, smoking and drinking habits, as well as

recreational substance use.

The family history (FH) involves gathering information about illnesses and deaths within the patient's immediate family. In this section, physicians routinely inquire about family occurrences of common problems such as cancer, heart disease, hypertension, diabetes, and other issues that may pose an increased risk in certain families. The systems review (SR) aims to extract additional information about the patient that might have been overlooked or omitted by other parts of the history. The SR can vary in length, being sometimes brief and other times quite extensive. The expectation, however, is that the physician will inquire about the patient's experience of symptoms in various organ systems, including but not limited to the skin, lymph nodes, head, eyes, ears, nose, throat, neck, etc. Following the physical examination (PE), the physician may initiate one or more additional investigations, such as lab tests, x-rays, electrocardiograms, etc. The intended purpose of these investigations is to clarify the diagnosis or gather data that may be useful for treatment or prevention.

6.1. Social Control in Medical Encounters

Social control in medical encounters refers to the ways in which societal norms, expectations, and regulations influence the behavior of individuals within the healthcare setting. Effective social control fosters trust in the healthcare system. Patients are more likely to seek medical help and follow prescribed treatments if they believe healthcare professionals adhere to ethical standards and are held accountable for their actions. Social control mechanisms, such as licensing boards and professional associations, hold healthcare professionals accountable for their conduct. This accountability is crucial for maintaining public trust and ensuring that practitioners meet certain standards of competence and ethical behavior. Assuring relevant search for specific medical information related to patients, and answering questions about some specific points associated with diagnosis making needs to be capable of taking into account social-level discourse features of medical encounter.

Through questioning, interruptions, and other means of redirecting the conversation, physicians selectively exclude certain topics from discussion while including others. Of particular focus are the verbal techniques employed to divert attention away from sources of personal distress in the social context. These techniques effectively prevent the critical consideration of the context and hinder the possibility of initiating change. The conveyance of ideologic messages and the invocation of social control in medical encounters are at times tied to physicians' explicit pronouncements on what patients should or should not do. Additionally, it is likely that ideology and social control emerge from the topics excluded from conversations between physicians and patients and how these exclusions come about.

Several studies on communication in medicine propose that medical encounters share common structural features. In a sociolinguistic examination of physician-patient conversations, West (1984) identifies typical "troubles" that arise when patients express concerns about events in their lives not readily addressed by physicians' technical intervention. The author argues that questions and interruptions serve as mechanisms by which physicians guide patients' concerns back to a technical track.

Understanding the connection between humanity and social control involves examining how societies establish and enforce rules, the impact of these mechanisms on individual freedom, and the ongoing negotiation between societal order and individual autonomy. It is an intricate interplay that evolves across cultures, contexts, and historical periods (Figure 16).



Figure 16. Humanity and social control.

6.2. Handling Voice of the Lifeworld

In his examination of medical encounters, Mishler (1984) illustrates how medical discourse severs connections to contextual issues and steers the focus toward technical matters. Mishler presents detailed transcripts derived from recordings of physician-patient communication (Waitzkin 1985) and outlines two conflicting "voices." The "voice of medicine" encompasses technical topics related to physiology, pathology, pharmacology, and other subjects relevant to doctors in their professional capacity. Conversely, the "*voice of the lifeworld*" encompasses the everyday, predominantly nontechnical problems that patients bring into the medical encounter.

According to Mishler's analysis of transcripts, patients frequently attempt to introduce contextual issues through the voice of the lifeworld. However, doctors may find themselves inadequately equipped to address such issues and, as a result, consistently revert to the voice of medicine. For instance, patients may bring up personal troubles unrelated to technical problems, or even if connected to technical issues, these personal troubles might not appear amenable to technical solutions. Alternatively, the introduction of personal troubles may lead to discomfort for the professional, the client, or both. In such situations, doctors commonly interject with questions, interruptions, or other tactics to shift the topic back to the voice of medicine.

6.3. Applications of Critical Discourse Analysis

Critical discourse analysis is the major theoretical and methodological framework that explores the social roles in relation to domination and social inequality. The personal challenges individuals bring to physicians often originate from societal issues extending beyond the realm of medicine. While medical encounters involve micro-level interactions, these individual processes unfold within a social context influenced by macro-level structures in society. This societal framework acts as a meta-language, governing relationships at the micro-level. A review of existing theories on medical discourse leads to the following propositions:

- (1) medical encounters often convey ideological messages supportive of the prevailing social order;
- (2) these encounters have implications for social control; and
- (3) medical language typically lacks a critical examination of the social context.

The technical structure of the medical encounter, as conventionally perceived by physicians, conceals a deeper structure that may not align closely with the conscious thoughts of professionals regarding their words and actions. Analogous patterns may emerge in interactions between clients and professionals in other helping professions. Whether expressed marginally or conveyed through an absence of criticism regarding contextual issues, ideology and social control in medical discourse

predominantly function as unintentional mechanisms aimed at securing consent.

Critical discourse analysis is formulated as an interdisciplinary and even transdisciplinary undertaking, acting as a response to the occasionally inflexible and rigid boundaries found within linguistics and other disciplines. This observation of interdisciplinarity manifests itself across three distinct levels: the theoretical foundations, the methodologies commonly employed by critical discourse scholars, and the research contexts in which critical discourse analysis is implemented (Unger 2016).

Fairclough's three-layered model of critical discourse analysis is shown in Figure 17. The 1st layer is inner, 2nd layer is middle, and 3rd layer is outer. (Ahmed et al 2017, Unger 2016). Critical discourse analysis can be viewed as a meta-discourse of object-level textual discourse with meta-relations of *interpretation, theory, conceptualization and operationalization* (Figure 17 on the right).

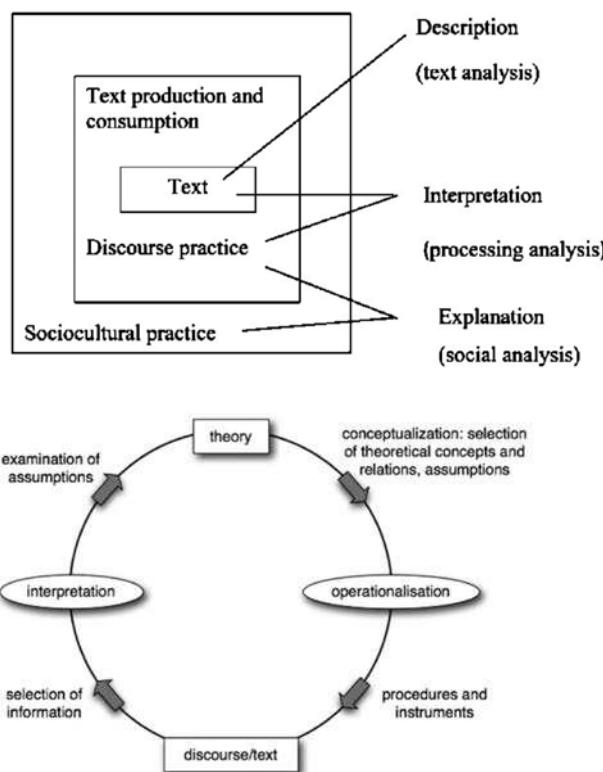


Figure 17. Visualization of critical discourse analysis.

6.4. Doctor-Patient Interaction Scenarios

A man visits his physician several months following a heart attack, expressing feelings of depression. With his disability payments set to end soon and his union on the verge of a strike, the physician informs him of his physical capability to resume work, emphasizing the positive impact of employment on his mental well-being. Additionally, the physician prescribes an antidepressant and a tranquilizer.

Figure 18 illustrates certain structural aspects of the discourse during the initial encounter. Viewed from this perspective, the contextual issue of uncertain employment becomes apparent (A). The patient experiences depression as a personal challenge in anticipation of returning to an uncertain job situation (B). Upon reaching the medical encounter (C), the patient tentatively and briefly voices concerns about an impending return to work coinciding with his union's plan to go on strike (D). Despite repeating these concerns at various points, the physician downplays their significance (E). Instead of delving into the contextual problem, the physician provides reassurance, emphasizing the positive impact of work on the patient's mental health. Additionally, the physician prescribes antidepressant medication and a tranquilizer (F). Post-encounter, it can be assumed that the patient continues preparing for his return to work.

A female patient consults a cardiologist due to irregularities in her heart rhythm, expressing concerns about palpitations and shortness of breath impacting her capacity to perform household chores. The physician conducts an electrocardiogram during her exercise, modifies her cardiac medications, and commends her for her dedication to upkeeping a well-organized household.

The structural components of the second encounter are depicted in Figure 19. In this context, the issue revolves around societal expectations regarding women's roles within the family:

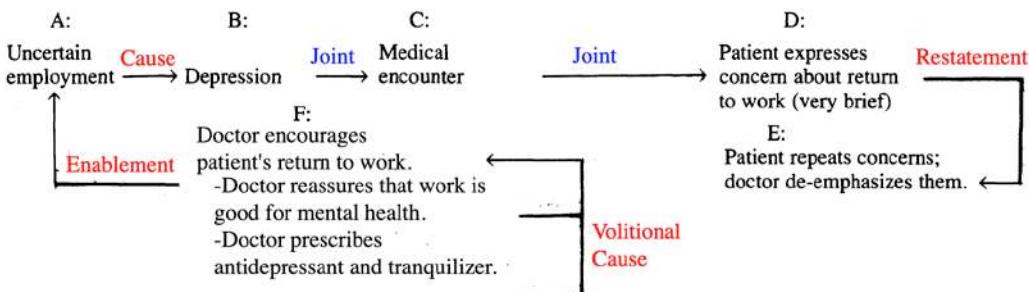


Figure 18. Structural elements of discourse and rhetorical relations between them (in red, significant; and blue, default).

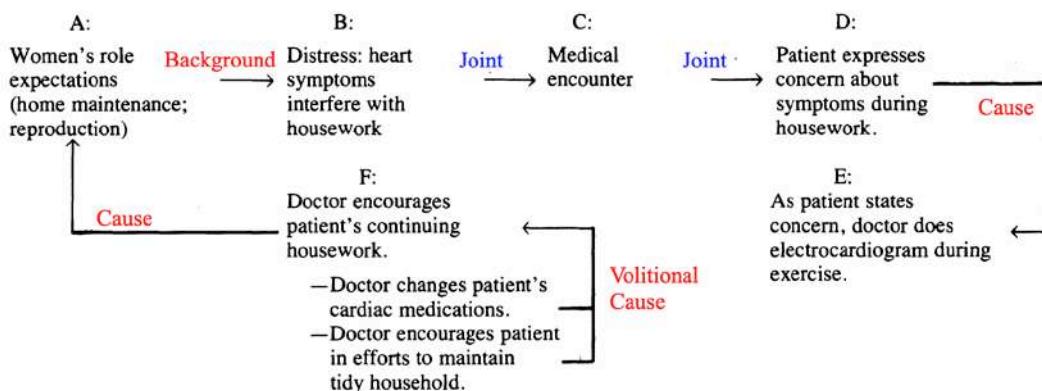


Figure 19. Structural Elements of medical encounter with a patient with heart symptoms.

(A) Housework, acknowledged by many as a crucial aspect of economic "reproduction," has traditionally fallen under the responsibility of women. The patient, whose heart symptoms impede her ability to perform housework, undergoes emotional distress as a result (B). Upon discussing this concern with her physician (C), the patient brings attention to this matter (D). Rather than delving deeply into her concern, the physician opts for an electrocardiogram while the patient engages in physical activity (E). Following the results, the physician adjusts the patient's cardiac medications and offers support for her efforts in maintaining a tidy household (F). Consequently, the patient grapples with the ongoing personal challenge of managing housework in the context of her severe heart disease.

Below there are samples of the data collected for reference purposes by (Adegbite and Odebunmi 2006)'s work:

Example 1.

Doctor: I was trying to find the veins of a year old baby to make intravenous infusion. Several unsuccessful attempts were made. The physician hissed in discouragement, and shook his hands several times in frustration.

Patients mother: (Down cast with red-eyeballs, yet held the baby tightly to the stretcher on which the physician was attending to her) Doctor, what exactly is the problem with my baby?

Doctor: She has acute respiratory infection but she will be okay.

Mother: (Broke down in tears and did not say a word)

This example illustrates a strong emotional distress which should not distract a QA system from answering a factual treatment-related question.

Example 2.

Doctor: Hello madam how is your health?

Patient: Doctor I am not well at all. I have been sick for more than 3 months but the problem is that I have been losing weight and I having persistent diarrhea. I have taken flagyl and chloraphenicol in vain.

Doctor: Do you cough?

Patient: Yes I do, but not so much.

Doctor: Do you have skin rash at the onset of this illness?

Patient: Yes, but it has disappeared.

Doctor: Any Fever?

Patient: No, but occasionally I feel as if I am having malaria.

Doctor: All right, before anything, we would have to do a series of tests on you. (Motioning the patient to stay outside)

Patient: (cuts in) Will I be admitted because I want to be under medical care in the hospital? Money is not my problem.

Doctor: Just go outside and relax. I will send you to the laboratory first before any treatment. (Doctor to the nurse) Take this card to the laboratory head. I have written some laboratory investigations that are to be carried out on the patient including '333' screening.

We apply health-specific labels to the discourse tree for this dialogue (Figure 20). The discourse tree encode the whole dialogue structure. Health-specific labels map the conversational structure into the logical flow of knowledge about the disease. [stage-medical-encounter] passes through: Chief complaint (CC), present illness (PI), past history (PH), and physical examination (PE). Communicative actions like 'send' are also important to encode the overall information flow. [social] helps to understand which dialogue utterances are not really related to health issues but a used to facilitate communication and proceed to the next stages of [stage-medical-encounter]. Mentioning fever indicates that [pain discourse] needs to be addressed.

elaboration

elaboration

joint

elaboration

TEXT:Doctor :

TEXT:Hello madam how is your health ?

elaboration

TEXT:Patient :

attribution (RightToLeft)

TEXT:Doctor

TEXT:I am not well at all . . . [stage-medical-encounter].

joint

elaboration

joint

TEXT:I have been sick for more than 3 months but the problem is that I have been loosing weight

TEXT:and I having persistent diarrhea .

joint

TEXT:I have taken flagyl and chloraphenicol in vain .

joint

TEXT:Doctor: Do you cough ?

contrast

TEXT:Patient: Yes I do ,

TEXT:but not so much .

joint

elaboration

TEXT:Doctor :

joint

attribution (RightToLeft)

TEXT:Do

TEXT:you have skin rash at the onset of this illness ?

elaboration

TEXT:Patient :

contrast (RightToLeft)

TEXT:Yes ,

TEXT:but it has disappeared .

joint

elaboration

TEXT:Doctor :

TEXT:Any Fever ?

contrast

TEXT:Patient : No , but occasionally I feel [pain discourse],

TEXT:as if I am having malaria .

elaboration

elaboration

TEXT:Doctor : All right , before anything , we would have to do a series of tests on you

[stage-medical-encounter],

topic-comment

attribution (RightToLeft)

TEXT:(Motioning the patient to stay outside) [social],

elaboration

same-unit

TEXT:Patient (cuts in) [patient-doctor dialogue structure],

TEXT:Will I be *admitted* because I want to be under medical care in the hospital ?

[communicative action]

elaboration

TEXT:Money is not my problem .
 elaboration
 TEXT:Doctor : Just go outside and relax .
 elaboration
 TEXT:I will *send* you to the laboratory first before any treatment . [stage-medical-encounter] [communicative action]
 TEXT:(Doctor to the nurse) Take this card to the laboratory head .
 elaboration
 TEXT:I have written some laboratory investigations
 elaboration
 TEXT:that are to be carried out on the patient
 TEXT:including ' 333 ' screening .

Figure 20. Discourse tree for complain from Example 2. We use the health-specific labels from Section 2.3.

We now proceed to an example of discourse tact and peacefulness in patient-doctor interactions. Example 3.

We conclude this subsection with examples of peaceful discourse (Figure 21) and discourse tact (Figure 22).

Patient's mother: She runs temperature every now and then... she's been unconscious now for three days.
 Doctor: [after examining baby] Yes, your daughter has cerebral malaria. The fever has got to some part of the brain. She has survival chances of 50–50.
 [Mother broke down in tears.] The disease cannot be cured, but we can try our best to control it.
 Patient's mother: Oh! God will help you.



Figure 21. Example of a peaceful discourse.

Example 4.

Patient: I hope what you are writing is not chloroquine?

 ChatGPT

Patient: I hope what you are writing is not chloroquine?

Doctor: I understand your concern, and I want to assure you that I take your health seriously. The medication I'm prescribing is not chloroquine. It's important for me to consider your well-being, and I'm here to answer any questions you may have about your treatment plan. Your health and safety are my top priorities, and I'm confident that the prescribed medication will contribute to your recovery. If you have any further concerns or if there's anything specific you'd like to discuss, please feel free to share, and we can address it together.

In this example, LLM produces most polite and cooperating but least informative reply Example 6.

Client: Doctor, he can't breathe very well.

Doctor: Yes he has acute respiratory problem.

Client: Is that why he can't breathe?

Doctor: Yes, we shall observe him for a while.



Figure 22. Example of a discourse tact: a caring adult and a stroller.

7. Patient-Doctor Communication Discourse

We now proceed from the general case of medical encounter to its particular case of patient-doctor dialogue. Recent linguistic research on patient-doctor communication falls into three main categories, as outlined by (Menz 2010):

- (1) Analyses focused on microstructure, examining conversational organization and interaction dynamics at the syntactic and semantic level.
- (2) Investigations exploring the impact of macrostructural social dimensions.
- (3) Practically-oriented studies assessing the social applicability of communication.

Research employing conversation analysis in patient-doctor communication is particularly concerned with various phases of this communication and the interactive tasks it serves. This type of research emphasizes formal and structural processes, with a growing interest in settings involving more than two participants. Additionally, there is an increasing focus on actual language usage and diverse forms of representing symptoms, disorders, and the subjective experience of illness in sociolinguistic analysis.

Research inspired by Conversational Analysis in the 1980s and 1990s utilized the frequency of interruptions or one-sided topic changes as indicators of practitioners' power dynamics. In a study focusing on gender in clinical interactions, patients could only initiate topic changes if physicians were agreeable, while doctors' topic changes seemed to require no such agreement. The study found

that male physicians made three times more unilateral topic changes than their female counterparts (Uskul & Ahmad 2003). Additionally, physicians might overlook expressed and embodied demonstrations of patients' suffering, impacting patient health but seemingly having no repercussions for the clinical relationship. In contrast, patients who ignore physicians may pose a threat to that relationship (Heath 1986). A comprehensive understanding of medical authority necessitates attention to discourse histories extending beyond individual clinical encounters, encompassing entire sequences of discursive interactions (Atkinson 1999), which may include discourse involving only care receivers or familial discussions.

The primary context for medical interaction is most comprehensively understood through the lens of the patient-doctor relationship (Fairclough 2003). This relationship serves as the cornerstone for building trust, establishing rapport, facilitating understanding, conveying diagnoses, and negotiating treatment. Consequently, the language used by physicians and patients plays a crucial role in shaping the comprehension of the patient's problem and molding the relationship itself, which can inherently possess a healing value.

Beran (1999) highlighted that patient-doctor interaction differs from ordinary or everyday spoken interaction due to its occurrence within an institutionalized setting. These institutions, particularly health care institutions, wield influence over all forms of discourse, and these discourses, in turn, are molded by broader power dynamics. This study systematically delves into the often intricate relationships of causality and determination existing between physicians and patients. It explores how such practices emerge from and are ideologically influenced by power relations and struggles over power.

Navigating the dynamic between a physician and a patient has become a contemporary challenge in our society. The evolution of this connection underscores the importance of informing patients and obtaining their consent for medical interventions. It becomes the responsibility of the physician to discern the information bound by medical confidentiality and what can be openly shared with the patient. A key objective of medical discourse is to foster a constructive dialogue between the physician and the patient, aiming to identify the root cause of the ailment, select a suitable treatment approach, and articulate actions using accessible vocabulary.

Within this conversation, medical specialists utilize specialized terminology and employ appropriate behavioral tactics. They are mindful that their communication can not only evoke positive emotions and reactions but also has the potential to trigger psychological trauma. Consequently, physicians bear the responsibility of delicately balancing these factors during their interactions with patients.

The medical practitioner engages in a conversation with a client with the aim of diagnosing the patient's problem. Simultaneously, the physician records notes on observations and prescriptions, forming a medical report intended for the client's treatment, stored in a dedicated medical file. The client may be a sick person, i.e., a patient, or the parent(s) or relation(s) of a sick person. The physician exerts control over the interaction by determining the pace of turn-taking (Adegbite 1991). They possess the authority to interrupt as needed and employ dominant acts such as directives, accusations, and caution to guide the client during the interaction. The success of diagnosis and treatment relies on the client's confidence in the medical system, a confidence built around the personality and care exhibited by the physician and other medical personnel (Adegbite and Odebunmi 2006).

7.1. Metaphorical Language

Discourse metaphor refers to a linguistic expression incorporating a construction that, in the relevant context, encourages the speaker/hearer to develop an analogical meaning negotiated within the discourse. This implies that discourse metaphors are specific to their form, as the analogy is triggered by a particular linguistic unit, namely a specific conventional form-meaning pairing. The discourse under consideration may range from a small group of speakers discussing a specific topic to all speakers engaging in mutually comprehensible utterances within a language community.

Explorations into metaphorical language within representations of illness represent a significant aspect of medical discourse, although the data in this area are not consistently drawn from natural contexts (Semino et al., 2004). Metaphors and other illustrative forms serve various purposes in the transfer of knowledge between experts and laypersons. Physicians may utilize these resources to explain complex facts, while patients may employ them to understand and convey sensations and experiences that are challenging to describe, such as the experience of pain or auras preceding epileptic seizures. Metaphors and similes are prevalent in these contexts, while exemplification and scenarios are more frequent when drawing parallels to everyday life (Figure 23). It is essential to note that these illustration processes are co-constructed, with experts and laypersons not necessarily using different resources but rather employing the same ones in distinct ways and for different purposes (Menz 2010).



Figure 23. Sometimes a health organization cannot fully keep the promise.

Analyzing variations in usage preferences can also contribute to the purposes of differential diagnostics. For instance, epilepsy patients tend to use metaphor more frequently in describing seizure attacks compared to patients with dissociative disorders. Significant differences also emerge in how each group reconstructs the gap in consciousness during attacks (Furchner, 2002). The analysis of disparities in linguistic strategies, therefore, proves useful in supporting differential diagnosis, traditionally a complex, costly, and error-prone process.

In a particular discourse, what are the discourse features that prompt the utilization of metaphorical language instead of opting for literal alternatives? For instance, what motivates individuals to express "grasp the essence" rather than "understand the meaning" in a specific context? Numerous NLP approaches to metaphorical language draw on cognitive and psycholinguistic insights, successfully formulating models for discourse coherence, abstractness, and affect. (Piccirilli and Schulte Im Walde 2022) establish cognitive and linguistic attributes such as frequency, abstractness, affect, discourse coherence, and contextualized word representations to anticipate the use of a metaphorical expression as opposed to a synonymous literal one in each context.

where $\{w_5, w_6\}$ are the two words composing the metaphorical expression, and $\{w_9, w_{10}\}$ are composing the literal paraphrase, which is shown in Figure 24. Depending on the expression input subject–verb or verb–object, the respective subject or object is identical in $\{w_5, w_6\}$ and $\{w_9, w_{10}\}$, as only the verb is used either as a metaphorical or a literal variant. The semantic relatedness between each word in $\{w_5, w_6\}$ and in $\{w_9, w_{10}\}$ is computed by the authors with each word in A.

Metaphorical language can be very diverse and

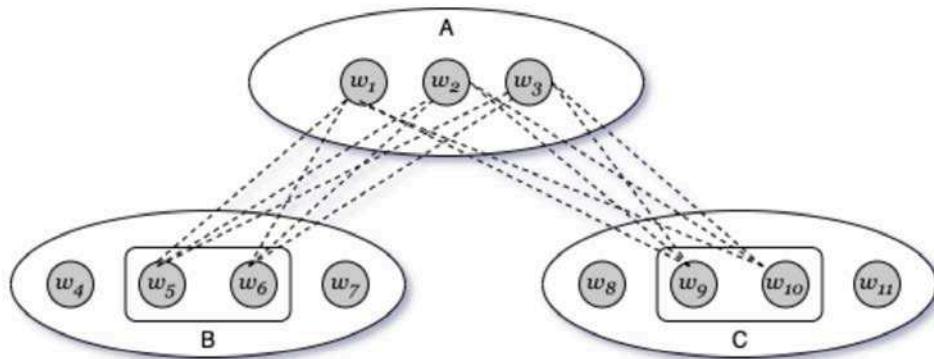


Figure 24. Illustration for a metaphoric discourse.

Let us have discourse A with three words $\{w_1, w_2, w_3\}$ and sentences B and C with four words,

7.2. Discourse of Pain Representation

The Cycle of Pain illustrates the problems that often happen when you live with pain. It is very common for one problem to lead to another, trapping you in a constant 'vicious cycle'. It can make you feel that things just continue to get worse and worse (Figure 25)

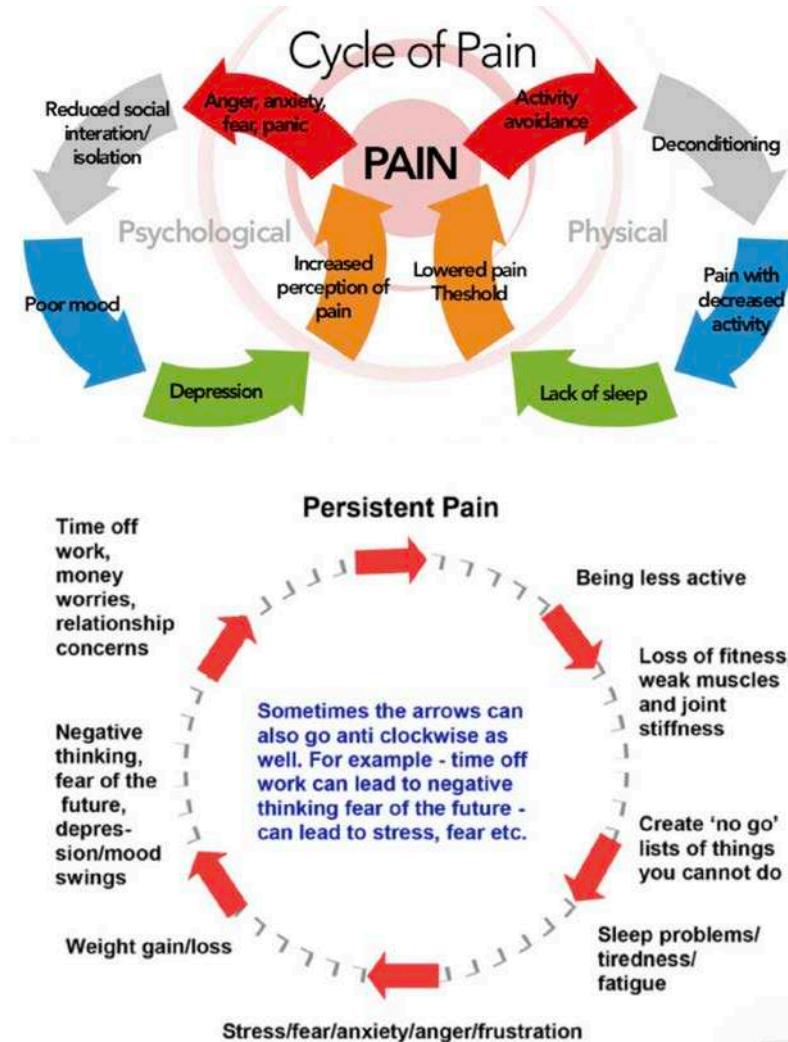


Figure 25. The cycle of pain (Moore 2023).

The portrayal of pain, a crucial aspect of medical communication, is often challenging due to the limited range of expression provided by everyday language. This limitation contrasts with, for example, the well-developed repertoires for describing visual or acoustic phenomena in most languages. From a medical perspective, category sets for representing pain have been established (Reisigl, 2009). These sets encompass dimensions such as temporal occurrence ('When?'), localization ('Where?'), intensity ('How severe?'), quality (e.g., 'stinging', 'piercing'), side symptoms (e.g., nausea), conditions of occurrence (e.g., when walking, when lying down), and pain management ('What eases or increases the pain?'). One challenge with such classifications is that it may not always be easy for patients to assign their subjective experience of pain to a specific medical category. Therefore, representations of pain often utilize non-verbal, gestural resources, which can be easily captured with recording technology and have become a significant focus in doctor-patient communication research.

The expression of the quality of pain is primarily verbal. The lack of 'basic pain terms' compels patients to use indirect means of description such as metaphor or visualization, as demonstrated in a recent study on written and oral German data (Overlach, 2008). In the oral context, lexical and syntactic variation was more strongly focused on basic metaphors of possession ('to have a pain') and copula construction ('the pain is ...').

When patients discuss their pain in non-medical settings, their conversation typically revolves around:

- (1) subjective theories regarding the illness and potential sources of the pain;
- (2) various impairments they experience due to the pain;
- (3) pain management strategies in general, including successful efforts to avoid pain or measures taken for relief.

As a result, the dimensions of conditions of occurrence and pain management take precedence. In contrast, in medical settings, the predominant themes include:

- (1) discussions about medication;
- (2) conversations about side symptoms associated with the pain that led to the medical consultation;
- (3) detailed specifications of the pain and its occurrence, covering the quality of the pain, as well as its local and temporal dimensions and intensity.

7.3. *Online Patient-Doctor Interactions*

Shang et al. (2019) explore the motivations and methods that lead patients to engage with physicians online (refer to Figure 26). A significant factor driving patients to opt for online consultations is the enhanced accessibility to physicians. The online platform provides a convenient space for patients to remotely connect with physicians, allowing them to discuss sensitive matters anonymously at their convenience without enduring lengthy waiting times and saving costs associated with hospital visits. Additionally, the desire for effective self-management of health conditions motivates patients to seek online consultations. Particularly, patients with long-term conditions, such as cancer, feel a responsibility for managing their conditions. This prompts them to independently search for solutions, empowering them to take control of preventive care, prepare for future consultations, and plan their treatment trajectory in physical hospitals. Moreover, patients turn to online counseling as a source of support due to unmet needs through offline channels. Dissatisfaction with previous offline healthcare experiences, marked by insufficient information, a lack of trust, and difficulty in understanding, prompts patients to seek a second opinion from online physicians.

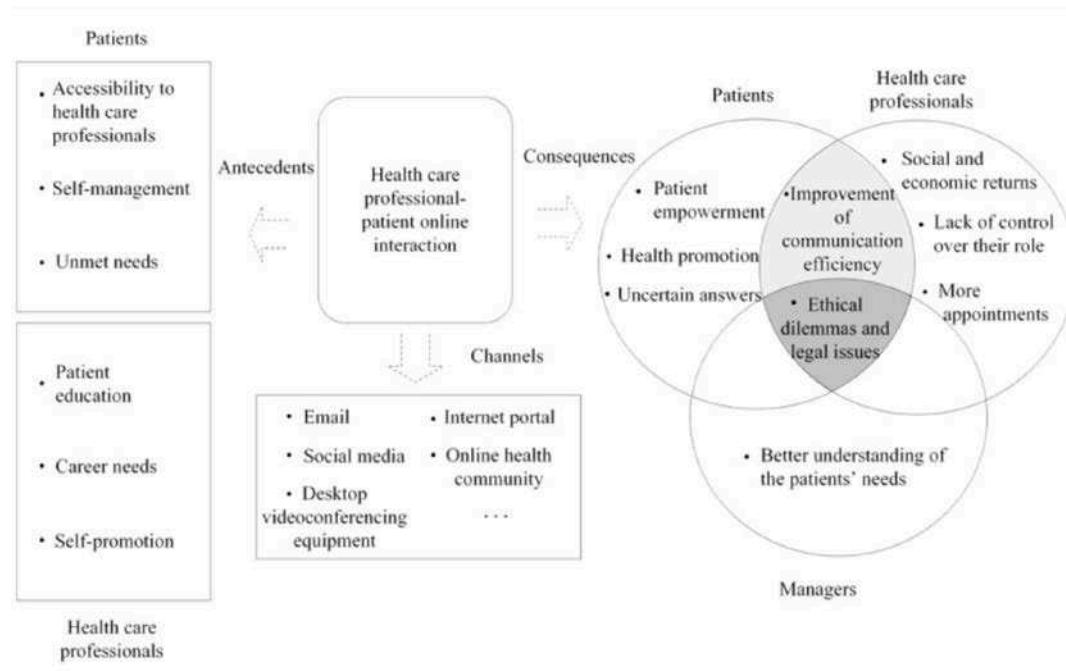


Figure 26. A higher-level view at physician-patient communication online.

7.4. Communicative Actions of Doctor-Patient Dialogue

Following the initial prefatory exchanges that involve greetings and summons, the transaction begins with the physician's initiation move, aiming to gather information about the nature and symptoms of the client's illness. This elicitation may reoccur in subsequent exchanges, occurring in opening, bound-opening, or re-opening moves. After the opening initiation, a response move follows, supporting it by providing a reply. If the reply is satisfactory, the physician proceeds with a follow-up supporting move, accepting the reply and proceeding to recommend prescriptions. However, if the reply is unsatisfactory, the physician may either re-open the elicitation or respond to the reply using pragmatic means to understand the problem, or both when necessary. The physician can employ challenging moves to condemn a patient's action, accuse or caution against excessive or incorrect behavior, or provide reassurance. Finally, the physician may use initiation moves to issue directives when recommending solutions to the client's problem.

Alternatively, a client may initiate a bound-opening move, offering more information to clarify a previous reply to the physician. The physician often supports the client by confirming these clarifications and assuring her that everything will be well. Occasionally, a client may check her understanding of a physician's suggestion, inviting a repetition of an earlier utterance, or request information from the physician, who then provides an answer.

7.4.1. Illocution

Illocution is a speaker's intended meaning, purpose, or communicative force behind the utterance.

Diagnosis as an institutional act in medicine is expressed in the conversation mainly via the general act of representatives. Representatives are, however, represented in individual utterance moves by such acts as elicitation, confirmation, comment, information, enquiry and conclusion.

The following acts can be identified with the participants in the interaction:

- (1) physician – *elicitation, explanation, confirmation, comment, assurance and criticize*
- (2) patient – *elicitation, complain/inform, request_explanation and appeal.*

The use of the directive act is ancillary. A physician uses it to caution or calm down a patient or to prepare him/her for medication and by the patient to appeal for pity or seek attention.

7.4.2. Perlocution

Perlocution refers to the impact or influence of a speaker's words on the listener or recipient. It is one of the three speech act categories introduced by philosopher Austin (1962), alongside locution and illocution. Locution pertains to the literal meaning of the words spoken, illocution refers to the intended or implied meaning, and perlocution focuses on the actual impact or response elicited in the listener. In simpler terms, perlocutionary acts involve the observable outcomes or consequences of speech on the audience, such as changes in attitudes, beliefs, emotions, or behaviors resulting from the communication. The effectiveness of communication can be assessed by examining the perlocutionary effects it produces.

In the realm of physician-patient diagnostic interaction, the utterances play a beneficial role for the participants. The client tends to be submissive to the dominance and control exerted by the physician in the interaction. Consequently, the directives and instructions provided by the physician are adhered to, and the physician's opinions are respected. Simultaneously, the physician pays careful attention to and is guided by the information shared by the client. Both the physician and the client collaborate in a joint effort to find solutions to problems, resulting in rare instances of argument or disagreement between them.

Politeness maxims and indirect communication acts are strategically employed to achieve a positive psychological impact on the patient. In the process of indirect communication, there are occasions when conversation maxims may be flouted, and pragmatic failure might unintentionally occur.

7.4.3. Locution of Utterances

Here is a brief description of the grammar of sentences in a typical physician-patient interaction from the perspective of systemic functional grammar (the grammatical terminologies are italicized). The interaction opens with an interrogative clause of the relational identifying type (are, is) in which the physician expresses a value (how? where?) of a token (your health). The client replies via a declarative clause of the relational attributive type (am, have been, has been) in which an attribute (not well, pregnant sick), is ascribed to a carrier (I, he, she). Alternatively, the reply is expressed via a declarative clause with the relational possessive process (have/has, am/is having got) in which a possessor (I, he, she) possesses items of illness - possessed (fever, malaria, cough, headache diarrhoea etc). Similar 'process' and 'participant' features to the ones above realize further diagnostic investigations in the interaction. Occasionally, however, there may be other clauses expressing either:

- (1) mental process of the reactional/affective type (feel[s]) in which a senser (I, he, she) is affected by a phenomenon or condition (hot, dizzy, like I'm having malaria); or
- (2) material process of the action type (eat, sleep, work or can't eat/sleep/work) in which participants are both the affected and goal in middle clauses.

In all of these expressions, circumstantial details of either inner or outer types may realize the time duration (3 months, for a long time); location (on my neck, in my mouth) and manner (persistently, seriously, properly, slowly) of an illness.

Patient: (Weak) Doctor, please, help me. I'm dying! Had I known I wouldn't have gone to him...

Doctor: Tell me what happened.

Patient: I got pregnant and because I'm not working yet and my partner too is yet to get a job so we decided to abort the pregnancy.

Doctor: How and where did you go for the termination?

Patient: I was taken to a physician's place and he used some instruments on me to remove the pregnancy.

Doctor: After that, what did you do again?

Patient: He gave me some drugs and injections. I'm still using those drugs.

Doctor: Do you have them here? Anyway, you have made a mistake of getting pregnant at first and for you to still go for termination in a wrong place—or what is the name of the clinic or hospital the operation was carried out?

Patient: It has no name. The physician treats people in his small apartment.

Doctor: Fake. You see, I hope the man has not punctured your womb, because with this fresh blood oozing out. (to her parents) she will definitely need blood. She's too pale and, not only that, she might ought to go surgical operation if bleeding persist.

Patient: What can I do I am in trouble. I pray God to forgive me and spare my life. Please help me.



ChatGPT

Doctor: I understand this is a difficult situation for you. First, let's focus on your well-being. It's crucial that we address the immediate concerns. I'll do everything in my power to help you. We need to assess the extent of any potential harm and stabilize your condition. You're not alone in this, and we'll work together to find the best course of action.

LLM can continue the dialogue in a meaningful, but uninformative way. However, it cannot provide an adequate action plan on physician's behalf:



You

what should doctor do, given the information from the patient:



ChatGPT

Given the serious situation described by the patient, the doctor should take the following steps:

1. Assess Vital Signs:

- Check the patient's vital signs, including heart rate, blood pressure, respiratory rate, and temperature. Assess the severity of the patient's condition and prioritize stabilization.

2. Address Immediate Medical Needs:

- If the patient is in critical condition, initiate life-saving measures such as administering oxygen, intravenous fluids, or blood transfusions if necessary.

3. Gather Detailed Medical History:

Regretfully, LLM provides a very broad, uninformative plan of action, which does not take into account a specific illness. A direct prompt needs to be formulated to obtain the information on what is expected from the physician:

**ChatGPT**

The specific treatment needed for the patient would depend on the extent of complications resulting from the unsafe abortion and the overall health of the individual. However, here are general considerations that a healthcare professional might take into account:

1. Stabilization:

- If the patient is in critical condition, the immediate focus will be on stabilizing vital signs. This may include administering intravenous fluids, oxygen therapy, and other measures to ensure the patient's physiological stability.

2. Assessment for Complications:

- Assess for complications related to the unsafe abortion, such as infection, hemorrhage, or damage to internal organs. This may involve imaging studies, blood tests, and a thorough physical examination.

3. Infection Management:

- If there are signs of infection, the patient may be prescribed antibiotics to treat or prevent the spread of infection. In severe cases, hospitalization may be necessary for intravenous antibiotic therapy.

Let us now see what is expected from the physician according to the human experts (Adegbite and Odebumi 2006)

In this situation, the physician should prioritize empathy, understanding, and reassurance to make the patient feel more comfortable. Here are some ways the physician can approach the dialogue:

1. Express Empathy:

- Doctor: I understand that you're going through a difficult time. I'm here to help you.

2. Reassure Confidentiality:

- Doctor: Your well-being is my priority, and everything you share with me will remain confidential.

3. Avoid Judgment:

- Doctor: I'm here to provide care, not to pass judgment. Let's focus on addressing your current health situation.

4. Encourage Open Communication:

- Doctor: It's important for me to know all the details so I can provide the best possible care. Please feel free to share everything about what happened.

5. Provide Comfort:

- Doctor: I'm here to support you, and together we will work towards a solution for your health concerns.

6. Avoid Blaming Language:

- Doctor: It's not uncommon for people to face challenging situations. Let's focus on moving forward and finding the best course of action for your health.

7. Address Immediate Concerns:

- Doctor: Right now, my main concern is ensuring your well-being. We'll discuss the details of what happened and work on a plan to address any health issues you're experiencing.

8. Explain Medical Steps:

- Doctor: I need to understand your medical situation better. Could you please share more details about the instruments used and the medications prescribed?

9. Discuss Potential Complications:

- Doctor: I'm concerned about the fresh blood and your overall condition. We need to address this promptly to ensure your safety. I might need to run some tests.

10. Reiterate Support:

- Doctor: You're not alone in this. We'll work together to find the best solution for your health. I'll involve the necessary medical interventions to ensure your well-being.

By approaching the conversation with empathy, understanding, and a focus on the patient's well-being, the physician can create a more comfortable environment for the patient to share information and receive appropriate medical care.

LLM also helps with additional questions the physician should ask:



You

Which questions should the doctor ask the patient to improve the treatment?



ChatGPT

To gather comprehensive information and improve the treatment plan, the doctor should ask the patient a series of questions covering various aspects of their health, medical history, and the circumstances surrounding the unsafe abortion. Here are some important questions the doctor might ask:

1. Current Symptoms:

- What specific symptoms are you experiencing right now?
- Can you describe the nature and intensity of any pain or discomfort?

2. Medical History:

- Do you have any pre-existing medical conditions?
- Are you currently taking any medications or undergoing any medical treatments?

3. Details of the Unsafe Abortion:

- Can you provide more details about the abortion procedure?
- What instruments were used, and were there any complications during the procedure?

7.5. From Medical Semantics to Sentiments and Discourse

Medical semantics is concerned with the meaning of medical terms and the relationships between them. It involves the study of how medical concepts are represented, interpreted, and related to one another. This field is more about the entities of medical knowledge, including their definition (such as diseases, treatments, and symptoms) and the relationships between these entities. Understanding that "hypertension" refers to high blood pressure and knowing how it is related to terms like "antihypertensive medications" in a systematic and structured way falls under medical semantics.

Medical discourse pertains to the communication, conversation, or written exchange that occurs within the medical field. It is about a though structure of parties involved in healthcare. It involves how healthcare professionals communicate with each other, with patients, and how medical information is conveyed in various contexts. The focus is on the actual use of language in medical contexts, including physician-patient interactions, medical writing, documentation, and communication within the healthcare team. An example of medical discourse is an analysis of how physicians discuss treatment options with patients, or examining the structure of medical reports and how they convey information.

Hence medical semantics deals with the underlying meaning and relationships of medical concepts, while medical discourse is concerned with the actual use of language and communication within the medical field. Both are important for effective communication and understanding in healthcare, with semantics providing the foundational knowledge structure, and discourse addressing the practical, contextual use of that knowledge.

(Heyn et al 2023) explore how the expression of positive emotions during the interaction between patients and providers can cultivate the patient-provider relationship. in the contexts of person orientation and positive outlook, patient-provider relationships improve by communication

conveying and eliciting positive emotions. The authors identify a number of underlying mechanisms which form either direct or indirect pathways between the context and the outcome.

Emotions represent affective, valenced responses to meaningful stimuli (Frijda, 2008), with positive emotions encompassing pleasant or desirable situational reactions. Research suggests that positive emotions contribute to the development of various cognitive resources, enhancing patients' life satisfaction, overall well-being, and functioning (Fredrickson, 2001).

The significance of expressions of positive emotions is underscored by the findings of Heyn et al. (2023), revealing direct and indirect pathways wherein communication strategies fostering positive emotions play a pivotal role in building and fortifying patient-nurse relationships. The role of positive emotional communication in cultivating patient-nurse relationships is depicted in Figure 27, where Contexts are linked to Mechanisms, leading to Outcomes. The edge weights signify the number of studies supporting a particular semantic relation. This semantic representation, connecting Contexts, Mechanisms, and Outcomes, can be interpreted as discourse on medical positiveness, applicable to both patient-nurse and patient-doctor dialogue notes.

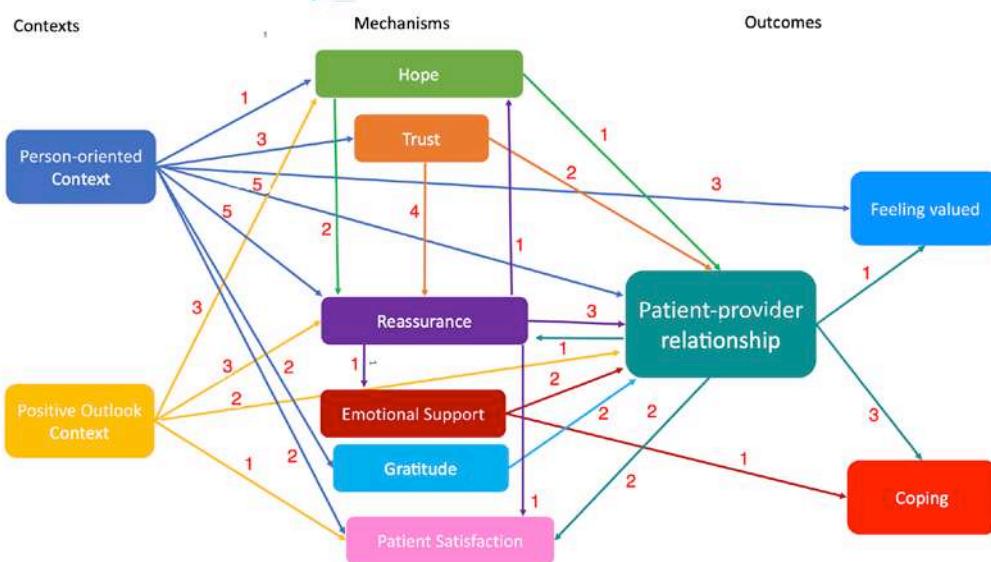


Figure 27. Positive sentiment discourse.

Patients articulated empathy as healthcare providers genuinely addressing their needs. When providers demonstrated care and understanding of the patients' situations, it led to a sense of validation for their emotions. Conversely, patients reported experiences with providers who lacked a person-oriented approach, often manifested non-verbally through tone of voice, body language, and a lack of presence. In such instances, patients felt marginalized, experienced a decline in hope, perceived suboptimal support (Bala et al., 2012), or developed a lack of trust in their providers (see Figure 28). Recognizing and responding to negative emotions emerges as a crucial aspect of effective healthcare communication. However, it underscores the need for a more balanced exploration of both positive and negative emotions in research and training related to communication in healthcare.

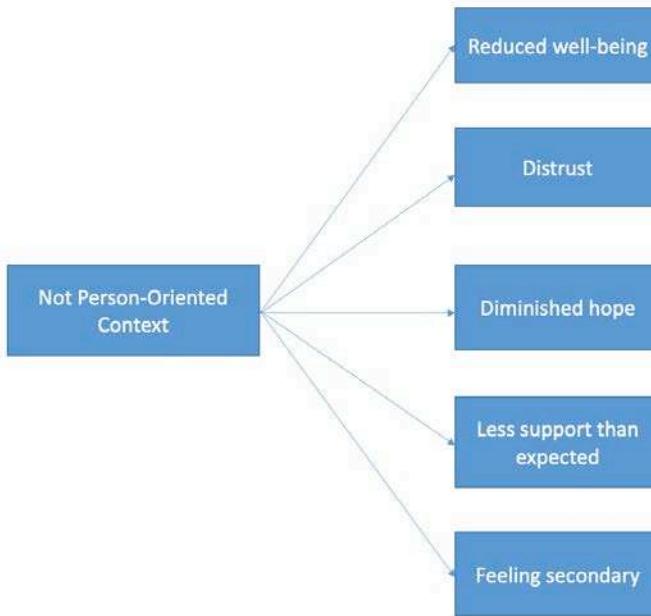


Figure 28. Negative sentiment discourse.

Galitsky and Kovalerchuk (2006) have devised an algorithm to aid in identifying email messages that may suggest individuals experiencing significant emotional distress. Numerous studies have established that terrorists often undergo substantial emotional distress before carrying out attacks. Therefore, the ability to detect emotional distress in individuals through email texts could potentially facilitate preventive measures. The proposed detection mechanism relies on extracting and classifying emotional profiles from emails. An emotional profile serves as a formal representation of a sequence of emotional states in a textual discourse, where communicative actions are linked to these emotional states. The authors associate an emotional profile with the classes "Emotional distress" or "No emotional distress," with the class assignment determined by an expert in a training dataset.

To illustrate, we present a snippet of correspondence between a prospective British suicide bomber (BBC 2005) and his relatives, who faced charges related to failing to notify authorities of a potential terrorist attack (see Figure 29). Identifying emotional distress in this context could have potentially averted a terrorist attack. On the left are selected fragments where emotions are highlighted in bold, and expressions amplifying them are in italic bold. On the right is an emotion intensity profile, ranging from negative to positive.

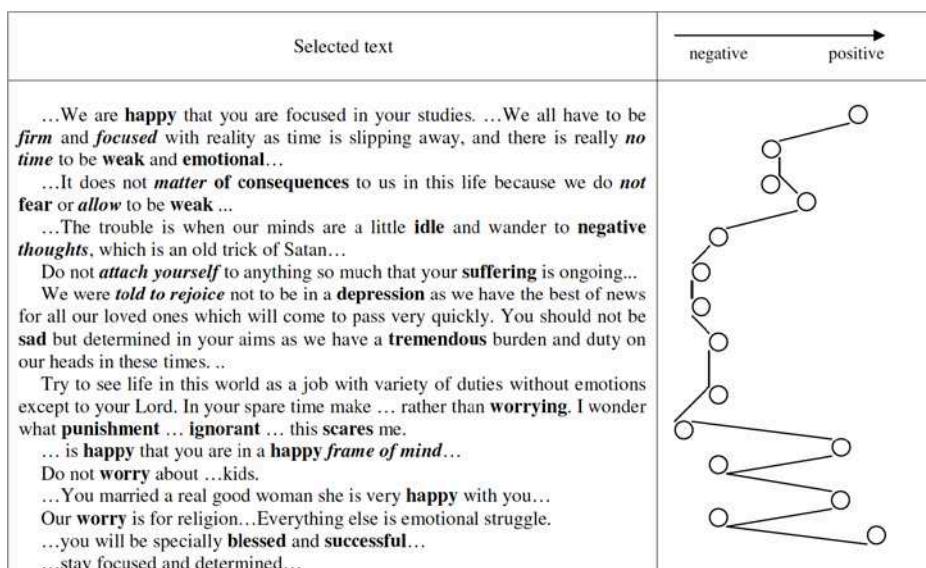
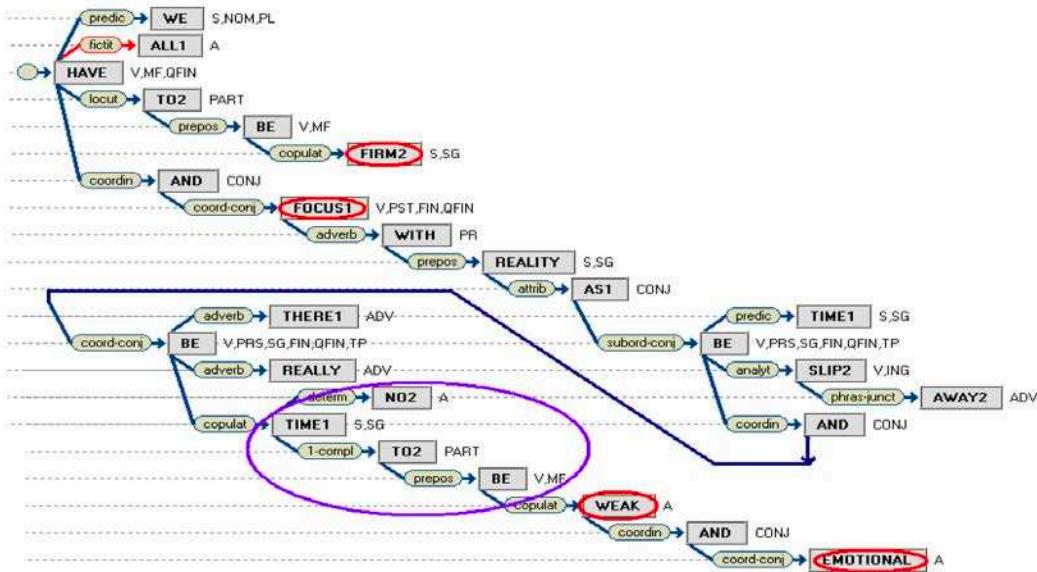


Figure 29. Emotional profile of a text.

There are multiple forms of expressions whose meanings can be classified as communicative actions or mental states; this example is a good illustration for how expressions indicating emotions are amplified. Also, one can see that a dependent occurrence of emotions amplifies their individual intensity (“*someone is happy that you are happy*”).

A parsing tree for the second sentence in Figure 29 is shown at Figure 30. Indications of emotions are shown in small ovals, we extract the words with explicit meanings for emotion (*firm, weak, emotional*) and the one which has a meaning of emotion because of the way it occurs in the sentence (*focus* in a passive voice). Emotions *weak, emotional* are amplified by the expression *no time to be* (shown by a larger oval) with the meaning “I encourage you to be”, which is an imperative communicative action.

**Figure 30.** A parse tree for a sentence .

8. Discourse and Retrieval Augmented Generation

Retrieval Augmented Generation (RAG) provides LLMs with the information retrieved from some data source to ground its generated answer on. RAG can be viewed as LLM prompting augmented with search, where one asks the LLM to answer the query provided the information found with the search algorithm as a context. Both the query and the retrieved context are injected into the prompt that is sent to the LLM (Ilin 2023).

RAG is the most popular architecture of the LLM based systems nowadays. There are many health systems built on top of RAG , from QA services combining web search engines with LLMs to apps with chat with specific health data. Vector similarity search area got stimulated by RAG.

Given a set of vectors x_i in dimension d , vector similarity search builds a data structure in memory from it. After the structure is constructed, when given a new vector x in dimension d , it

performs efficiently the operation:

where $\|\cdot\|$ is the Euclidean distance (L^2).

Hence, the data structure is an *index*, an object that has an *add* method to add x_i vectors. Note that the x_i 's are assumed to be fixed. Computing the *argmin* is the search operation on the index (Figure 31).

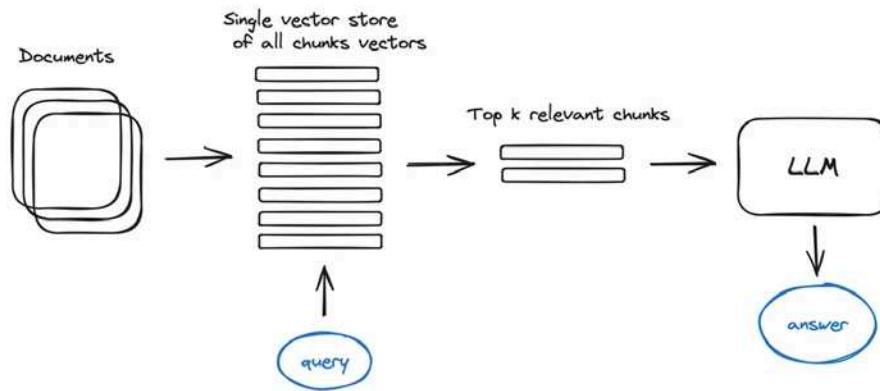


Figure 31. Index search.

Simple RAG case is as follows:

- (1) texts is split into chunks,
- (2) these chunks are embedded into vectors with some Transformer Encoder model,
- (3) all those vectors are saved into an index
- (4) a prompt is created for an LLM that requests the model to answers user's query given the context identified in the search step.

In the runtime:

- (1) the user's query is embedded with the same Encoder model
- (2) the search is executed of this query vector against the index, find the top-k results,
- (3) the corresponding text chunks are retrieved from our database
- (4) these chunks are fed into the LLM prompt as context.

An index of vectors is created, representing the document contents so that at runtime we search for the least cosine distance between all these vectors and the query vector which corresponds to the closest semantic meaning. Chunking and vectorization is required. Default chunking is splitting the initial documents in chunks of some size without loosing their meaning according to sentence or paragraph boundaries.

The crucial part of the RAG pipeline is the search index, storing the vectorized content of a document. The most naive implementation uses a just flat index which is a brute force distance calculation between the query vector and all the chunks' vectors. A proper search index, optimized for efficient retrieval on million+ elements scales is a vector index like Faiss, using some approximate Nearest Neighbor implementation (Chap ??) like clustering or trees (Figure 32).

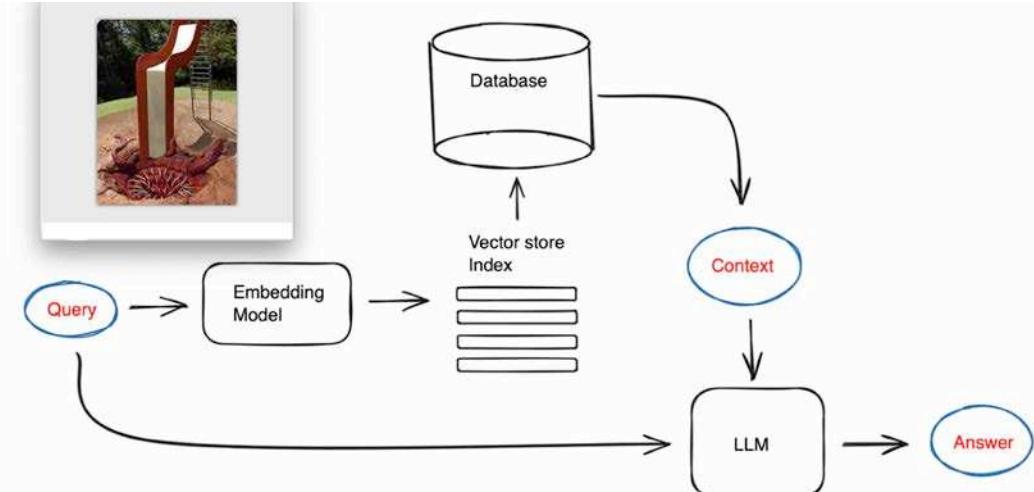


Figure 32. A simple RAG architecture. People search input is shown on top-left.

We also use fusion retrieval which takes the best from both worlds:

- (1) keyword-based old school search; sparse retrieval algorithms like TF*IDF or search industry standard BM25, and
- (2) modern semantic or vector search,

combining these in one retrieval result. The only trick here is to properly combine the retrieved results with different similarity scores; this problem is usually solved with the help of the Reciprocal Rank Fusion algorithm, reranking the retrieved results for the final output.

Reciprocal Rank Fusion sorts the documents according to a naive scoring formula. Given a set D of documents to be ranked and a set of rankings R , each a permutation on $1..|D|$, the RRF score is computed (Cormack et al 2009):

$$RRFscore(d \in D) = \sum_{r \in R} \frac{1}{k + r(d)},$$

The intuition in choosing this formula derived from fact that while highly-ranked documents are more important, the importance of lower-ranked documents does not vanish as it would were, say, an exponential function used. The constant k mitigates the impact of high rankings by outlier systems

A RAG architecture leveraging discourse analysis is shown in Figure 33. We build a discourse representation of a long, complex query. We also build a hierarchical representation of chunks of documents in the form of discourse tree as an addition index, along with Vector store index and summary index.

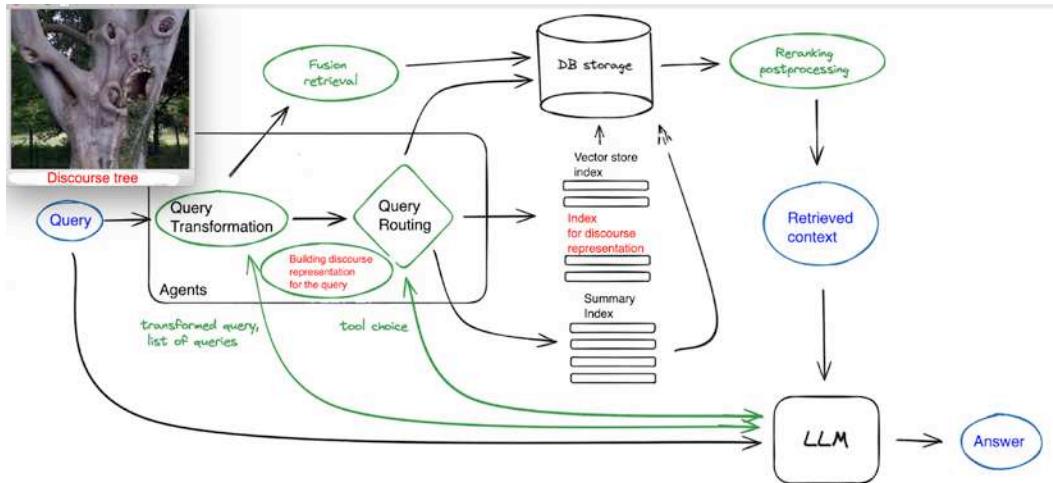


Figure 33. RAG architecture with discourse analysis.

9. Conclusions

We developed the *MedDiscourse* system to answer queries against both unstructured and structured medical documents such as electronic health record, doctor's notes and patient-doctor conversation logs relying on LLMs.

Building the system, the medical discourse literature was consulted; we came up with application of discourse features for responding to inquiries within lengthy and intricate medical documents. Our investigation included a thorough examination of dialogue logs with patients; we developed a discourse model specifically oriented for the health domain.

Within our discourse model, we integrated the structure of patient interviews, tackled the metaphoric language spoken by patients and health professionals, addressed various patient-doctor communication modalities found in text, and implemented a specialized discourse mechanism for online patient-doctor dialogues. We expanded the conventional notion of a discourse tree, covering the whole document. Proposed extension of the discourse tree concept covers the broad spectrum of medical writing styles.

It was confirmed that effective discourse analysis requires an understanding of the social context in patient-doctor interactions to filter out preliminary answers which are affected by social norms instead of correct health information. Taking into account the unique aspects of online doctor-patient communication, including intentions, motivations, career advancement considerations and patients' trust, we proposed our approach to provide adequate answers and identify the root causes of illnesses and other problems and issues. It turned out that discourse features can reveal concealed or implicit data during the diagnostic process. This is expected to handle cases with missing information in text. We observed that ascending to the discourse level can compensate a lack of common sense and medical knowledge needed for QA that relies upon a deep understanding of lengthy health documents with diverse structures and styles.

Our proposed approach adopts a neuro-symbolic paradigm, where the LLM serves as the baseline for question-answering, and discourse analysis operates at the symbolic level, effectively "spreading" question-answering capabilities across lengthy, unstructured documents.

In case there are many documents to retrieve from, one need to efficiently search inside them, find relevant information and synthesize it in a single answer with references to the sources. An efficient way to do that in case of a large database is to create two indices:

- (1) one composed of summaries and
- (2) the other one composed of document chunks,

and to search in two steps, first filtering out the relevant docs by summaries and then searching just inside this relevant group.

Employing data and structure-rich representation is expected to increase QA accuracy. Ablation study results are depicted in Figure 34

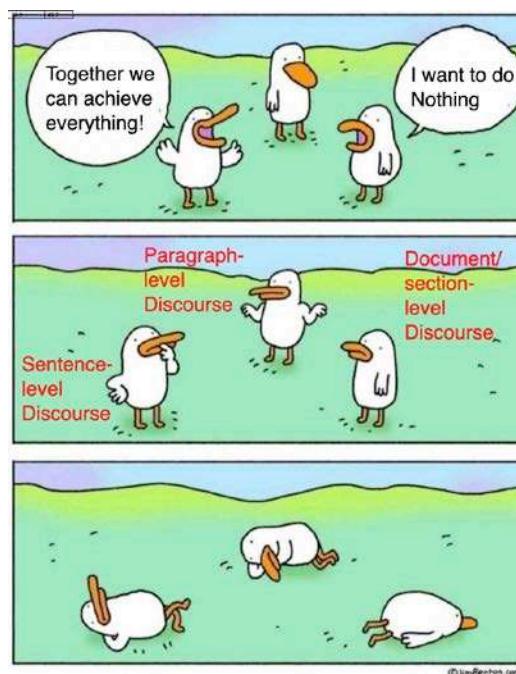


Figure 34. Combining discourse representations at different levels.

References

1. Afantinos S, Eric Kow Nicholas Asher Jérémie Perret (2015) Discourse parsing for multi-party chat dialogues. In EMNLP, pp 928–937
2. Ahmed, Zahid & Su, Ling & Ahmed, Khalid. (2017). Pakistani youth manipulated through night-packages advertising discourse. Human Systems Management. 36. 151-162. 10.3233/HSM-171762.
3. Ainslie J, Santiago Ontanon, Chris Alberti, Vaclav Cvcek, Zachary Fisher, Philip Pham, Anirudh Ravula, Sumit Sanghai, Qifan Wang, and Li Yang. 2020. Etc: Encoding long and structured inputs in transformers. In EMNLP, pages 268–284.
4. American Psychologist, 56(3), 218–226.

5. Atkinson P (1999) Medical discourse evidentiality and the construction of professional responsibility. In: Talk, Work and Institutional Order, De Gruyter Mouton <https://doi.org/10.1515/9783110208375.2.75>
6. Austin, J.L. 1962. How to do things with words. London: Oxford University Press.
7. Bala, S. V., Samuelson, K., Hagell, P., Svensson, B., Fridlund, B., & Hesselgard, K. (2012). The experience of care at nurse-led rheumatology clinics. *Musculoskeletal Care*, 10(4), 202–211.
8. BBC (2005) Suicide bomber trial: Emails in full. http://news.bbc.co.uk/2/hi/uk_news/3825765.stm
9. Beltagy I, Matthew E. Peters, Arman Cohan (2020) Longformer: The Long-Document Transformer. arXiv:2004.05150
10. Beran, J. (1999). Doctor-Patient Communication: Part I- Introduction. Prague, Czech Republic: Karolinum .
11. Cao S and Lu Wang. 2022. HIBRIDIS: Attention with Hierarchical Biases for Structure-aware Long Document Summarization. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 786–807, Dublin, Ireland
12. Chen, Bodong & Zhu, Xinran & Shui, Hong. (2022). Socio-Semantic Network Motifs Framework for Discourse Analysis. [10.1145/3506860.3506893](https://doi.org/10.1145/3506860.3506893).
13. Chi MTH and Ruth Wylie. 2014. The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational psychologist* 49, 4 (Oct. 2014), 219–243. <https://doi.org/10.1080/00461520.2014.965823>
14. Cohan A, Franck Dernoncourt, Doo Soon Kim, Trung Bui, Seokhwan Kim, Walter Chang, and Nazli Goharian. 2018. A discourse-aware attention model for abstractive summarization of long documents. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers), pages 615–621, New Orleans, Louisiana.
15. Cordella, M. 2004. The dynamic consultation: a discourse analytical study of physician-patient communication. Amsterdam: John Benjamins.
16. Dasigi P, Kyle Lo, Iz Beltagy, Arman Cohan, Noah A. Smith, and Matt Gardner. 2021. A dataset of information-seeking questions and answers anchored in research papers. In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 4599–4610
17. Du H, Y Feng, C Li, Y Li, Y Lan, D Zhao (2023) Structure-Discourse Hierarchical Graph for Conditional Question Answering on Long Documents. Findings of the Association for Computational Linguistics: ACL, 6282-6293
18. Fairclough, N. (2003). Analyzing Discourse. London and New York: Routledge.
19. Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions.
20. Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist*, 56(3), 218–226.
21. Frijda, N. H. (2008). The psychologists'point of view. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (Vol. 3, pp. 68–88). The Guilford Press.
22. Furchner, I. (2002) ' "Keine absence gleicht der anderen." Die Darstellung von Bewusstseinslücken in "Anfallsbeschreibungen"', in G. Brünner and E. Gülich (eds), *Krankheit verstehen. Interdisziplinäre Beiträge zur Sprache in Krankheitsdarstellungen*. Bielefeld: Aisthesis Verlag. pp. 121–42
23. Galitsky B (2017) Matching parse thicket for open domain question answering. *Data & Knowledge Engineering* 107, 24-50
24. Galitsky B (2019a) Learning Discourse-Level Structures for Question Answering. In *Developing Enterprise Chatbots*. Springer, Cham Switzerland. 177-219
25. Galitsky B and Kuznetsov SO (2008) Learning communicative actions of conflicting human agents. *J Exp Theor Artif Intell* 20(4):277–317
26. Galitsky B, Boris Kovalerchuk (2006) Mining emotional profiles using e-mail messages for earlier warnings of potential terrorist activities *Proceedings Volume 6241, Data Mining, Intrusion Detection, Information Assurance, and Data Networks Security*.
27. Gao Y, Chien-Sheng Wu, Jingjing Li, Shafiq Joty, Steven CH Hoi, Caiming Xiong, Irwin King, and Michael Lyu. 2020. Discern: Discourse-aware entailment reasoning network for conversational machine reading. *EMNLP*, pages 2439–2449.
28. Girnth, H. (1996) Texte im politischen Diskurs. Ein Vorschlag zur diskursorientierten Beschreibung von Textsorten'. *Mutterspache* 1 66-80.
29. Heath C. 1986. Body Movement and Speech in Medical Interaction. Cambridge, UK: Cambridge Univ. Press
30. Heyn LG, Løkkeberg ST, Ellington L, van Dulmen S, Eide H. Understanding the role of positive emotions in healthcare communication - A realist review. *Nurs Open*. 2023 Jun;10(6):3447-3459.

31. Ilin I (2023) Advanced RAG Techniques: an Illustrated Overview. Towards AI. <https://pub.towardsai.net/advanced-rag-techniques-an-illustrated-overview-04d193d8fec6>
32. Izacard G and Édouard Grave. 2021. Leveraging passage retrieval with generative models for open domain question answering. In Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume, pages 874–880
33. Joty S, Giuseppe Carenini, Raymond T. Ng, and Gabriel Murray (2019) Discourse Analysis and Its Applications. ACL Tutorial Abstracts, pages 12–17 Florence, Italy, July 28 - August 2
34. Joty SR, Carenini G, Ng RT, Mehdad Y (2013) Combining intra-and multi- sentential rhetorical parsing for document-level discourse analysis. In: ACL, vol. 1, pp 486–496
35. Kipper K, Korhonen A, Ryant N, Palmer M (2008) A large-scale classification of English verbs. Language Resources and Evaluation Journal 42:21–40
36. Labov, W. and J. Waletzky, 'Narrative Analysis: Oral Versions of Personal Narratives', in: J. Helm (ed.). Essays on the Verbal and Visual Arts: Proceedings of the 1966 Annual Spring Meeting of the American Ethnological Society. (Seattle: University of Washington Press, 1967), pp. 12-44.
37. Mann, W.C., Thompson, S.A., 1988. Rhetorical structure theory: towards a functional theory of text organization. Text 8 (3), 243–281.
38. Menz F (2010) Doctor-Patient Communication. In book: "The Sage Handbook of Sociolinguistics" Editors: Johnstone B, Wodak R, Kerswill P. Sage.
39. Milo R, Shalev Itzkovitz, Nadav Kashtan, Reuven Levitt, Shai Shen-Orr, Inbal Ayzenshtat, Michal Sheffer, and Uri Alon. 2004. Superfamilies of Evolved and Designed Networks. Science 303, 5663 (March 2004), 1538–1542. <https://doi.org/10.1126/science.1089167>
40. Mishler, Elliot G. 1984. The Discourse of Medicine: Dialectics of Medical Interviews. Norwood, N.J. : Ablex.
41. Moore P (2023) The pain toolkit . <https://www.sheffieldachesandpains.com/assets/professional-PDFs-2/persistent/Pain%20Toolkit.pdf>
42. Nair, I., Somasundaram, S., Saxena, A., & Goswami, K. (2023). Drilling Down into the Discourse Structure with LLMs for Long Document Question Answering. arXiv:2311.13565
43. Nie Y, Heyan Huang, Wei Wei, and Xian-Ling Mao. 2022. Capturing global structural information in long document question answering with compressive graph selector network. In Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, pages 5036–5047, Abu Dhabi, United Arab Emirates.
44. PatientInfo (2023) Heavy leg lifting after 3 months. <https://patient.info/forums/discuss/heavy-leg-lifting-after-3-months-805587>
45. Piccirilli P and Sabine Schulte Im Walde. 2022. What Drives the Use of Metaphorical Language? Negative Insights from Abstractness, Affect, Discourse Coherence and Contextualized Word Representations. In Proceedings of the 11th Joint Conference on Lexical and Computational Semantics, pages 299–310, Seattle, Washington.
46. Prasad R, Dinesh N, Lee A, Miltsakaki E, Robaldo L, Joshi A, Webber B (2008) The Penn Discourse Treebank 2.0. In: Proceedings of the 6th international conference on language resources and evaluation (LREC), Marrakech, Morocco
47. Reisigl, Martin, Wodak, Ruth and Meyer, M. (2009). Methods of critical discourse analysis: Introducing qualitative methods. In "Methods of Critical Discourse Analysis". Edited by:Ruth Wodak & Michael Meyer. SAGE Publications, Ltd
48. Ribeiro, B.T. 2002. Erotic strategies in powerless women: analyzing psychiatric interviews, in Gender identity and discourse analysis. Sunderland, J. & Litosseliti, L. (eds.). Amsterdam: John Benjamins, pp. 193–219.
49. Saad-Falcon J, Joe Barrow, Alexa Siu, Ani Nenkova, David Seunghyun Yoon, Ryan A. Rossi, Franck Dernoncourt (2023) PDFTriage: Question Answering over Long, Structured Documents. arXiv:2309.08872
50. Schauer H 2000. From elementary discourse units to complex ones. In 1st SIGDial Workshop on Discourse and Dialogue, pages 46–55.
51. Schiffrin, D. Approaches to Discourse. Language as Social Interaction. (Cambridge, MA: Blackwell, 1994).
52. Semino, E., Heywood, J. and Short, M. (2004) 'Methodological problems in the analysis of metaphors in a corpus of conversations about cancer', Journal of Pragmatics: an Interdisciplinary Journal of Language Studies, 36: 1271–94.
53. Shang L, Zuo M, Ma D, Yu Q. The Antecedents and Consequences of Health Care Professional-Patient Online Interactions: Systematic Review. J Med Internet Res. 2019 Sep 25;21(9):e13940. doi: 10.2196/13940.

54. Stahl G and Kai Hakkarainen. 2021. Theories of CSCL. In International Handbook of Computer-Supported Collaborative Learning, Ulrike Cress, Jun Oshima, Carolyn Rosé, and Alyssa Wise (Eds.). Springer, Berlin.
55. Sun H, William Cohen, and Ruslan Salakhutdinov. 2022. ConditionalQA: A complex reading comprehension dataset with conditional answers. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 3627–3637.
56. Sun H, William W Cohen, and Ruslan Salakhutdinov. 2021. End-to-end multihop retrieval for compositional question answering over long documents.
57. Taboada M and William C. Mann. 2006. Rhetorical Structure Theory: Looking Back and Moving Ahead. *Discourse Studies*, 8(3):423–459, June.
58. Unger, J. 2016. The interdisciplinarity of critical discourse studies research. *Palgrave Commun* 2, 15037.
59. Uskul AK, Ahmad F. 2003. Physician–patient interaction: a gynecology clinic in Turkey. *Soc. Sci. Med.* 57:205–15
60. van Dijk, Teun. (2009) *Critical Discourse Studies: A Sociocognitive Approach* 1. Methods of Critical Discourse Analysis.
61. Waitzkin H. 1985. Information Giving in Medical Care." *Journal of Health and Social Behavior* 26:81-101.
62. West, Candace. 1984. *Routine Complications: Troubles With Talk Between Doctors and Patients*. Bloomington: Indiana University Press.
63. Wittgenstein, L. *Philosophische Untersuchungen*. (Frankfurt am Main: Suhrkamp, 1967).
64. Yu N, Meishan Zhang, Guohong Fu, and Min Zhang. 2022. RST discourse parsing with second-stage EDU level pre-training. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 4269–4280
65. Zengin, Mevlüde. (2016). An Introduction to Intertextuality as a Literary Theory: Definitions, Axioms and the Originators. *Pamukkale University Journal of Social Sciences Institute*. 2016. 299-327. 10.5505/pausbed.2016.96729.
66. Zummo, Marianna (2012). In-Between Discourse and Genre: Doctor-Patient Interaction in Online Communication.. *Romanian Journal of English Studies*. 9. 10.2478/v10319-012-0009-8.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.