

Sequencing Legal DNA

NLP for Law and Political Economy

12. Discourse and Knowledge

What is the endpoint of NLP?

What is the endpoint of NLP?

Machine understanding of text **discourse**.

What is the endpoint of NLP?

Machine understanding of text **discourse**.

- ▶ good summaries of long text: extraction of relevant information, discarding of irrelevant information.

What is the endpoint of NLP?

Machine understanding of text **discourse**.

- ▶ good summaries of long text: extraction of relevant information, discarding of irrelevant information.
- ▶ knowledge graphs: structured representations of knowledge learned from unstructured texts.

What is the endpoint of NLP?

Machine understanding of text **discourse**.

- ▶ good summaries of long text: extraction of relevant information, discarding of irrelevant information.
- ▶ knowledge graphs: structured representations of knowledge learned from unstructured texts.
- ▶ question answering: reading questions and summarizing the relevant answers from the knowledge graph.

What is the endpoint of NLP?

Machine understanding of text **discourse**.

- ▶ good summaries of long text: extraction of relevant information, discarding of irrelevant information.
- ▶ knowledge graphs: structured representations of knowledge learned from unstructured texts.
- ▶ question answering: reading questions and summarizing the relevant answers from the knowledge graph.
- ▶ storytelling: enriched text generation that understands characters and their relationships.

Outline

Discourse

Text Summarization

Knowledge Graphs

Question Answering

Dialogue Systems

Tools for Legal NLP

Discourse

- The coherent structure of language above the level of sentences or clauses.
 - ▶ Coherence: meaningful connections between the items in the sequence.
 - ▶ Applications: coreference resolution, automated essay grading, automatic summarization, dialogue systems

Source for this topic: Jurafsky and Martin book, Yulia Tsvetkov slides.

Co-Reference Resolution

“My sister has a cat. Her name is Roberta.”



[Cat's] name is Roberta \leftrightarrow [Sister's] name is Roberta

Coherence

“My sister has a cat. Her name is Roberta.”



“My sister has a cat. She attended a conference.”

Coherence

“My sister has a cat. Her name is Roberta.”



“My sister has a cat. She attended a conference.”

- ▶ Coherent sentences are connected by structured relations. Some examples from a standard schema:
 - ▶ **Reason:** Sentence 2 gives a reason for the action in Sentence 1.
 - ▶ **Elaboration:** Sentence 2 provides additional information or detail about the situation presented in Sentence 1
 - ▶ **Evidence:** Sentence 2 provides information to convince the reader to accept Sentence 1.
 - ▶ **Attribution:** Sentence 2 gives the source of information in Sentence 1.

Coherence

“My sister has a cat. Her name is Roberta.”



“My sister has a cat. She attended a conference.”

- ▶ Coherent sentences are connected by structured relations. Some examples from a standard schema:
 - ▶ **Reason:** Sentence 2 gives a reason for the action in Sentence 1.
 - ▶ **Elaboration:** Sentence 2 provides additional information or detail about the situation presented in Sentence 1
 - ▶ **Evidence:** Sentence 2 provides information to convince the reader to accept Sentence 1.
 - ▶ **Attribution:** Sentence 2 gives the source of information in Sentence 1.
- ▶ Another perspective: Coherent sentences are centered around the same entities.

Learning Discourse Relations

- ▶ Xu et al (2019) use a model like skip-thought embeddings to produce coherence relations between adjacent sentences:
 - ▶ embed sentences and predict whether they are adjacent
 - ▶ trained model gives a coherence score as the probability of being adjacent

Learning Discourse Relations

- ▶ Xu et al (2019) use a model like skip-thought embeddings to produce coherence relations between adjacent sentences:
 - ▶ embed sentences and predict whether they are adjacent
 - ▶ trained model gives a coherence score as the probability of being adjacent
- ▶ Nie et al (2019) use “discourse markers” (and, but, because, etc) to build a huge labeled dataset of discourse relations between sentences.
 - ▶ they fine-tune BERT embeddings to predict discourse relations, producing “DisSent” embeddings.

Outline

Discourse

Text Summarization

Knowledge Graphs

Question Answering

Dialogue Systems

Tools for Legal NLP

Text Summarization

Goal: produce a shorter version of a text that contains the most relevant or important information.

Text Summarization

Goal: produce a shorter version of a text that contains the most relevant or important information.

- ▶ obvious applications in law / legal practice.

Text Summarization

Goal: produce a shorter version of a text that contains the most relevant or important information.

- ▶ obvious applications in law / legal practice.
- ▶ could also work as a dimension reduction or information extraction for social science measurements.

Single Document Summarization

Document

Cambodian leader Hun Sen on Friday rejected opposition parties ' demands for talks outside the country , accusing them of trying to " internationalize " the political crisis .

Government and opposition parties have asked King Norodom Sihanouk to host a summit meeting after a series of post-election negotiations between the two opposition groups and Hun Sen 's party to form a new government failed .

Opposition leaders Prince Norodom Ranariddh and Sam Rainsy , citing Hun Sen 's threats to arrest opposition figures after two alleged attempts on his life , said they could not negotiate freely in Cambodia and called for talks at Sihanouk 's residence in Beijing .Hun Sen , however , rejected that ."

I would like to make it clear that all meetings related to Cambodian affairs must be conducted in the Kingdom of Cambodia , " Hun Sen told reporters after a Cabinet meeting on Friday ." No-one should internationalize Cambodian affairs .

It is detrimental to the sovereignty of Cambodia , " he said .Hun Sen 's Cambodian People 's Party won 64 of the 122 parliamentary seats in July 's elections , short of the two-thirds majority needed to form a government on its own .Ranariddh and Sam Rainsy have charged that Hun Sen 's victory in the elections was achieved through widespread fraud .They have demanded a thorough investigation into their election complaints as a precondition for their cooperation in getting the national assembly moving and a new government formed



Summary

Cambodian government rejects opposition's call for talks abroad

Summarization for question answering: Snippets

- Google: 156 characters (about 26 words) plus title and link

The screenshot shows a Google search results page. At the top left is the Google logo. To its right is a search bar containing the query "what is die brücke?". Below the search bar, the word "Search" is highlighted in red. To its right, it says "About 5,910,000 results (0.28 seconds)". A horizontal line separates this from the search results. On the left side of the results, there is a sidebar with categories: "Everything", "Images", "Maps", "Videos", "News", "Shopping", "Applications", and "More". Each category has a corresponding link and a brief snippet of text. The "Everything" section shows the most relevant results.

Category	Result Title	Result Description
Everything	Die Brücke - Wikipedia, the free encyclopedia	en.wikipedia.org/wiki/Die_Br%C3%BCcke Die Brücke (The Bridge) was a group of German expressionist artists formed in Dresden in 1905, after which the Brücke Museum in Berlin was named. Founding ... You've visited this page 5 times. Last visit: 4/16/12
News	Die Brücke (film) - Wikipedia, the free encyclopedia	en.wikipedia.org/wiki/Die_Br%C3%BCcke_(film) Die Brücke (English: The Bridge) is a 1959 West German film directed by Austrian filmmaker Bernhard Wicki. It is based on the eponymous 1958 novel by ...
More	Die Brücke - Die Brücke Art	www.huntfor.com/arthistory/c20th/diebrucke.htm Die Brücke was the association of artist expressionists from Dresden, Germany. ... Die Brücke made use of a technique that was controlled, intentionally ...

Source: Chan Young Park & Dan Jurafsky Slides

- ▶ Extractive summarization:
 - ▶ create the summary from phrases or sentences in the source document(s)
- ▶ Abstractive summarization:
 - ▶ express the ideas in the source documents using (at least in part) different words

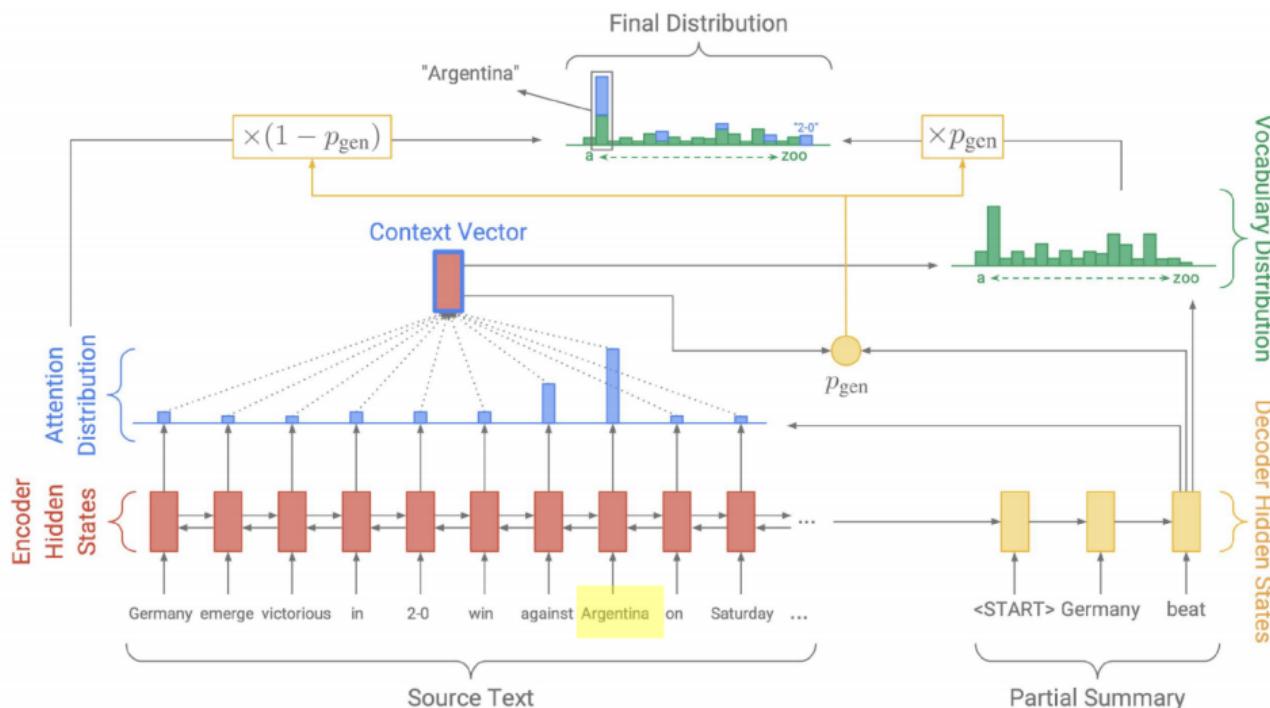
- ▶ Extractive summarization:
 - ▶ create the summary from phrases or sentences in the source document(s)
- ▶ Abstractive summarization:
 - ▶ express the ideas in the source documents using (at least in part) different words
- ▶ A tough-to-beat extractive baseline:
 - ▶ take the first sentence.

Breaking summarization into steps:

1. content selection: choose sentences to extract from the document
 - ▶ e.g., sentences with words with high IDF weights
2. information ordering: choose an order to place them in the summary
3. sentence realization: clean up the sentences

Pointer-Generator Networks (See, Liu, and Manning 2017)

Copy words from the source text



Pointer-Generator Networks

Article: andy murray (...) is into the semi-finals of the miami open , but not before getting a scare from 21 year-old austrian dominic thiem, who pushed him to 4-4 in the second set before going down 3-6 6-4, 6-1 in an hour and three quarters. (...)

Summary: andy murray **defeated** dominic thiem 3-6 6-4, 6-1 in an hour and three quarters.

Article: (...) wayne rooney smashes home during manchester united 's 3-1 win over aston villa on saturday. (...)

Summary: manchester united **beat** aston villa 3-1 at old trafford on saturday.

Figure 7: Examples of abstractive summaries produced by our model (**bold** denotes novel words).

TLDR: Summarization of Scientific Documents

Paper

Abstract: We present the Compressive Transformer, an attentive sequence model which compresses past memories for long-range sequence learning. We find the ...
Introduction: Humans have a remarkable ability to remember information over long time horizons. When reading a book,...

tl;dr (author)

Long-range transformer using a compressive memory, achieves sota in wikitext-103 and enwik8 LM benchmarks, release a new book-level LM benchmark PG-19.

Peer reviews

This paper presents a new variation of the Transformer model, named Compressive Transformer. The key novelty of this model is to preserve long range memory in a compressed form, instead of discarding them as previous models have done

tl;dr (readers)

A new Transformer model that preserves long memory in compressed form and improves over character LM benchmarks.

Figure 1: A TLDR is an extreme summary of a scientific paper. Our SCITLDR dataset includes multiple TLDRs for each paper. One written by the authors of papers and others are manually rewritten from peer reviews.

Outline

Discourse

Text Summarization

Knowledge Graphs

Question Answering

Dialogue Systems

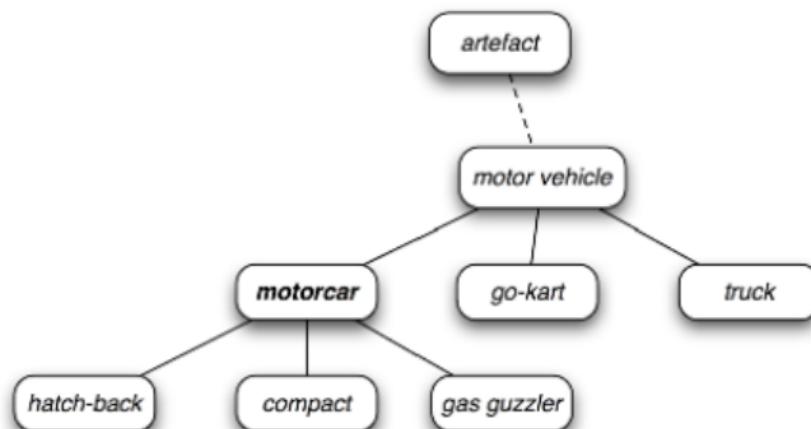
Tools for Legal NLP

Knowledge Graph

- ▶ A structured graph representing facts (and assertions?) as tuples.
- ▶ Entities are nodes, relations are edges:
 - ▶ (head entity, relation, tail entity)

Knowledge Graph

- ▶ A structured graph representing facts (and assertions?) as tuples.
- ▶ Entities are nodes, relations are edges:
 - ▶ (head entity, relation, tail entity)
- ▶ E.g., WordNet:



- ▶ A crowd-sourced effort to extract structured information from Wikipedia and make it available as linked open data.

GENERATING FACTS FOR THE ENTITY BILLIE HOLIDAY

“Facts” as RDF Triples



Subject Predicate Object
(Thing)



S <http://dbpedia.org/resource/Billie_Holiday>
P <<http://xmlns.com/foaf/0.1/name>>
O "Billie Holiday"

KNOWLEDGE BASE CONSTRUCTION PROJECTS

Method	Schema	Examples
Curated	Yes	Cyc/OpenCyc [23], WordNet [24], UMLS [25]
Collaborative	Yes	Wikidata [26], Freebase [7]
Auto. Semi-Structured	Yes	YAGO [4, 27], DBPedia [5], Freebase [7]
Auto. Unstructured	Yes	Knowledge Vault [28], NELL [6], PATTY [29], PROSPERA [30], DeepDive/Elementary [31]
Auto. Unstructured	No	ReVerb [32], OLLIE [33], PRISMATIC [34]

Information Extraction

- ▶ Manual information extraction is too costly:
 - ▶ e.g., in FreeBase knowledge graph, 71% of humans are missing date of birth.
- ▶ Can use relation extraction from text to fill in the graph.

Information Extraction

- ▶ Manual information extraction is too costly:
 - ▶ e.g., in FreeBase knowledge graph, 71% of humans are missing date of birth.
- ▶ Can use relation extraction from text to fill in the graph.
- ▶ e.g. "United has a hub in Chicago, which is the headquarters of United Continental Holdings":
 - ▶ {United; has a hub in; Chicago}
 - ▶ {Chicago; is the headquarters of; United Continental Holdings}

Information Extraction

- ▶ Manual information extraction is too costly:
 - ▶ e.g., in FreeBase knowledge graph, 71% of humans are missing date of birth.
- ▶ Can use relation extraction from text to fill in the graph.
- ▶ e.g. "United has a hub in Chicago, which is the headquarters of United Continental Holdings":
 - ▶ {United; has a hub in; Chicago}
 - ▶ {Chicago; is the headquarters of; United Continental Holdings}
- ▶ Standard rule-based approach uses dependency parse to extract relations (Bank et al 2007, Fader et al 2011):
 - ▶ filters: relation must contain a predicate; subject and object must be noun phrases.
 - ▶ thresholds: aggregate over large corpora and keep relatively frequent (therefore potentially reliable) relations.

Named Entity Recognition

Type	Tag	Sample Categories	Example sentences
People	PER	people, characters	Turing is a giant of computer science.
Organization	ORG	companies, sports teams	The IPCC warned about the cyclone.
Location	LOC	regions, mountains, seas	The Mt. Sanitas loop is in Sunshine Canyon.
Geo-Political Entity	GPE	countries, states, provinces	Palo Alto is raising the fees for parking.
Facility	FAC	bridges, buildings, airports	Consider the Golden Gate Bridge.
Vehicles	VEH	planes, trains, automobiles	It was a classic Ford Falcon.

Figure 18.1 A list of generic named entity types with the kinds of entities they refer to.

[PER Washington] was born into slavery on the farm of James Burroughs.
[ORG Washington] went up 2 games to 1 in the four-game series.
Blair arrived in [LOC Washington] for what may well be his last state visit.
In June, [GPE Washington] passed a primary seatbelt law.
The [VEH Washington] had proved to be a leaky ship, every passage I made...

Figure 18.3 Examples of type ambiguities in the use of the name *Washington*.

- Blackstone has a trained legal NER system in spaCy (for UK law).

Soares et al, Unsupervised Learning of Relations

r_A	In 1976, e_1 (then of Bell Labs) published e_2 , the first of his books on programming inspired by the Unix operating system.
r_B	The “ e_2 ” series spread the essence of “C/Unix thinking” with makeovers for Fortran and Pascal. e_1 ’s Ratfor was eventually put in the public domain.
r_C	e_1 worked at Bell Labs alongside e_3 creators Ken Thompson and Dennis Ritchie.
Mentions	e_1 = Brian Kernighan, e_2 = Software Tools, e_3 = Unix

Table 2: Example of “matching the blanks” automatically generated training data. Statement pairs r_A and r_B form a positive example since they share resolution of two entities. Statement pairs r_A and r_C as well as r_B and r_C form strong negative pairs since they share one entity in common but contain other non-matching entities.

- ▶ Soares et al (2019) use Google Natural Language API to extract entities and assign unique ID's.
 - ▶ could also use spaCy NER, for example.
- ▶ Train BERT with special entity marker tokens before and after entity spans.

Learning from Knowledge Graph Relations

- ▶ Knowledge graph relations can be predicted from other relations:
 - ▶ e.g., parents of a person are often married, so (John, parent of, Anne) + (Mary, parent of, Anne) → (John, married to, Mary)

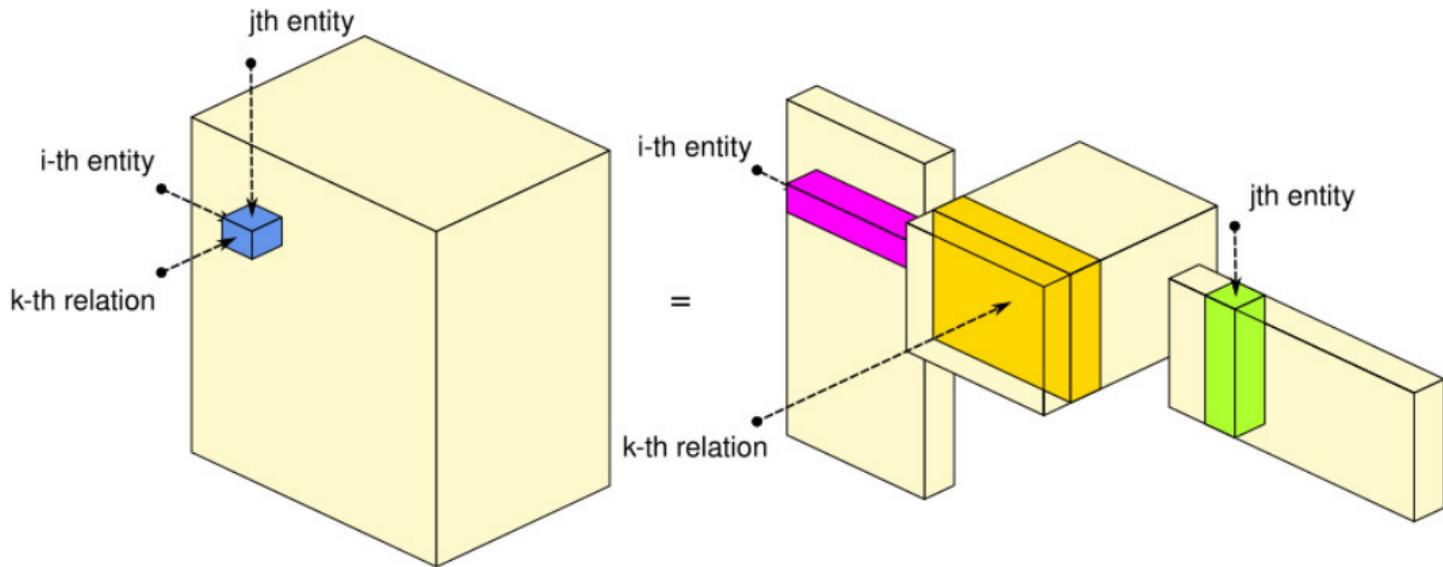
Learning from Knowledge Graph Relations

- ▶ Knowledge graph relations can be predicted from other relations:
 - ▶ e.g., parents of a person are often married, so (John, parent of, Anne) + (Mary, parent of, Anne) → (John, married to, Mary)

EXAMPLES OF PATHS LEARNED BY PRA ON FREEBASE TO PREDICT WHICH COLLEGE A PERSON ATTENDED

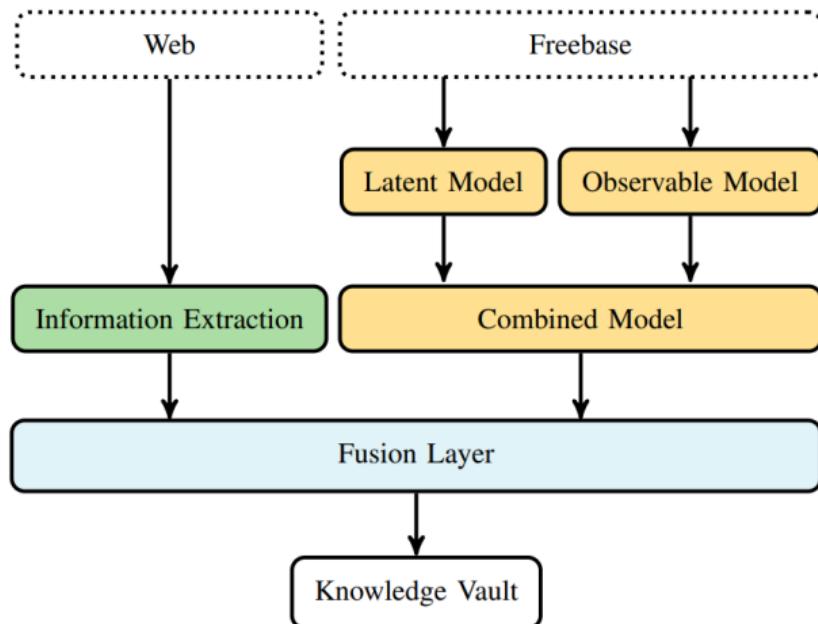
Relation Path	F1	Prec	Rec	Weight
(draftedBy, school)	0.03	1.0	0.01	2.62
(sibling(s), sibling, education, institution)	0.05	0.55	0.02	1.88
(spouse(s), spouse, education, institution)	0.06	0.41	0.02	1.87
(parents, education, institution)	0.04	0.29	0.02	1.37
(children, education, institution)	0.05	0.21	0.02	1.85
(placeOfBirth, peopleBornHere, education)	0.13	0.1	0.38	6.4
(type, instance, education, institution)	0.05	0.04	0.34	1.74
(profession, peopleWithProf., edu., inst.)	0.04	0.03	0.33	2.19

Tensor Factorization for Knowledge Graph Embeddings



- ▶ RESCAL: three-way factorization method for (entity, relation, entity) tuples.
 - ▶ quadratic computational costs in embedding dimension, so doesn't scale well.
 - ▶ Nickel et al (2015) use circular convolutions to help this method scale.
- ▶ Joulin et al (2017) show state-of-the-art results for predicting edges with fasttext.

Knowledge Vault Architecture



Using structured knowledge in language models: Hayashi et al (2019)

Topic: **Barack Obama**

Article **Barack Hussein Obama II** (...; born August 4, 1961) is an American[nationality] attorney[occupation] and politician[occupation] who served as the 44th president of the United States[position held] from 2009 to 2017. ...

Knowledge Graph

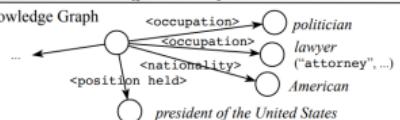


Figure 1: Overview of our task of language modeling conditioned on structured knowledge. For a given topic, we want to learn an LM that leverages the knowledge graph through relations when modeling the text.

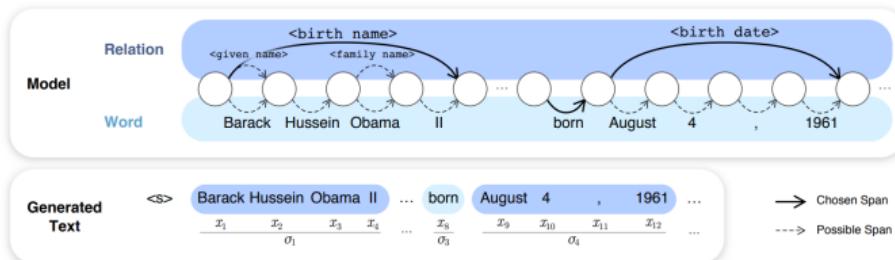


Figure 2: While generating, our model switches between the two sources, namely “Relation” and “Word”. Nodes represent hidden states up to each token, and edges represent possible span matches, i.e., choice of latent variables. In this example, we show one choice of latent variables with solid lines, and other options as dashed lines. We also show an “annotation” of the token sequence by the spans and sources we choose.

Outline

Discourse

Text Summarization

Knowledge Graphs

Question Answering

Dialogue Systems

Tools for Legal NLP

Information Retrieval for Question Answering

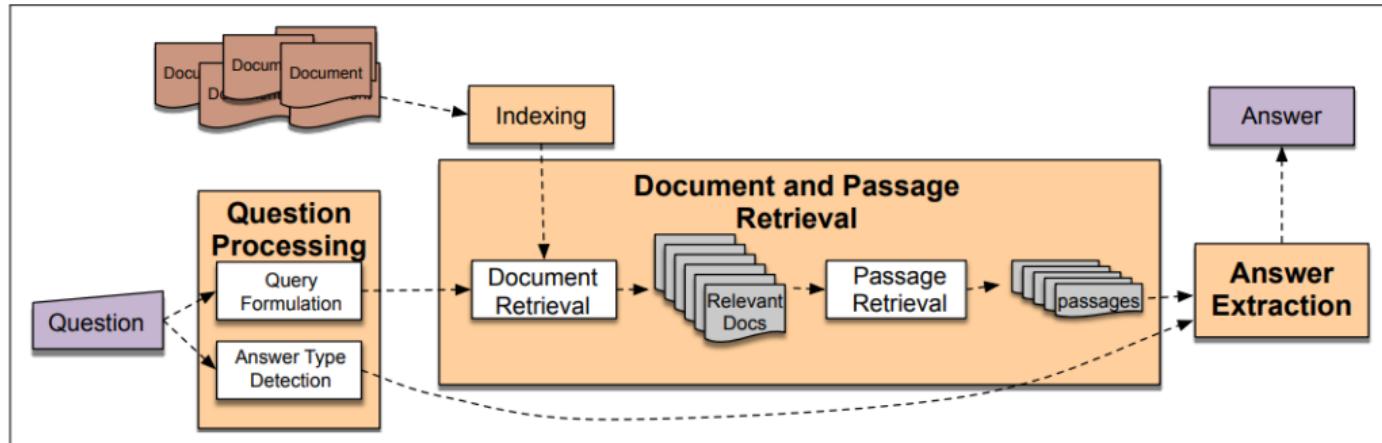


Figure 25.2 IR-based factoid question answering has three stages: question processing, passage retrieval, and answer processing.

- ▶ e.g., IBM Watson is a fast search engine over a knowledge base.

SQuAD Question Answering Dataset

Beyoncé Giselle Knowles-Carter (born September 4, 1981) is an American singer, songwriter, record producer and actress. Born and raised in [Houston, Texas](#), she performed in various [singing and dancing](#) competitions as a child, and rose to fame in the late 1990s as lead singer of R&B girl-group Destiny's Child. Managed by her father, Mathew Knowles, the group became one of the world's best-selling girl groups of all time. Their hiatus saw the release of Beyoncé's debut album, [Dangerously in Love \(2003\)](#), which established her as a solo artist worldwide, earned five Grammy Awards and featured the Billboard Hot 100 number-one singles "Crazy in Love" and "Baby Boy".

Q: "In what city and state did Beyoncé grow up?"

A: "[Houston, Texas](#)"

Q: "What areas did Beyoncé compete in when she was growing up?"

A: "[singing and dancing](#)"

Q: "When did Beyoncé release [Dangerously in Love](#)?"

A: "[2003](#)"

Figure 25.6 A (Wikipedia) passage from the SQuAD 2.0 dataset (Rajpurkar et al., 2018) with 3 sample questions and the labeled answer spans.

- ▶ BERT variants currently getting state-of-the-art performance on SQuAD.

Pre-trained models from huggingface

```
from transformers import pipeline
qapipe = pipeline('question-answering')
qapipe({
    'question': """how can question answering service produce
answers""",
    'context': """One such task is reading comprehension. Given a
passage of text, we can ask questions about the passage that can be
answered by referencing short excerpts from the text. For instance,
if we were to ask about this paragraph, "how can a question be
answered in a reading comprehension task" ..."""
})
```

output:

```
{'score': 0.38941961529900837,
'start': 128,
'end': 169,
'answer': 'referencing short excerpts from the text.'}
```

Outline

Discourse

Text Summarization

Knowledge Graphs

Question Answering

Dialogue Systems

Tools for Legal NLP

Chatbots

- ▶ Chatbots have been around since the 1960s, and a system from 1968 already passed the Turing test.

Chatbots

- ▶ Chatbots have been around since the 1960s, and a system from 1968 already passed the Turing test.
- ▶ Rule-based systems use pattern-action rules.
- ▶ Corpus-based systems use information retrieval + encoder-decoder networks

ELIZA: Rule-Based Therapy Chatbot

```
function ELIZA GENERATOR(user sentence) returns response
    Find the word w in sentence that has the highest keyword rank
    if w exists
        Choose the highest ranked rule r for w that matches sentence
        response  $\leftarrow$  Apply the transform in r to sentence
        if w = 'my'
            future  $\leftarrow$  Apply a transformation from the 'memory' rule list to sentence
            Push future onto memory stack
        else (no keyword applies)
            either
                response  $\leftarrow$  Apply the transform for the NONE keyword to sentence
            or
                response  $\leftarrow$  Pop the top response from the memory stack
    return(response)
```

Neural Chatbots

- ▶ Think of response generation as a task of transducing from the user's prior turn to the system's turn.
- ▶ Train on:
 - ▶ movie dialogue databases
 - ▶ Twitter conversations
- ▶ Train a deep neural network
 - ▶ map from user1 turn to user2 response

Neural Chatbots

- ▶ Think of response generation as a task of transducing from the user's prior turn to the system's turn.
- ▶ Train on:
 - ▶ movie dialogue databases
 - ▶ Twitter conversations
- ▶ Train a deep neural network
 - ▶ map from user1 turn to user2 response
- ▶ Better performance with adversarial nets and reinforcement learning.

Meena (Google Chatbot, 2019)

Human: I like how cute cows are

Meena: Ah, interesting. You like how they are cute, right?

Human: Yeah, and they're really smart.

Meena: I heard they go to college

Human: Cows go to college?

Meena: I heard that a cow went to Harvard.

Human: What did the cow study?

Meena: Bovine sciences.

Human: Do horses go to Harvard?

Meena: Horses go to Hayvard.

Human: That's a pretty good joke, I feel like you led me into it

Meena: You were trying to steer it elsewhere, I can see it

Ethical issues in dialog systems

- ▶ Machine learning systems replicate biases that occurred in the training data.
- ▶ Microsoft's Tay chatbot:
 - ▶ Went live on Twitter in 2016, taken offline 16 hours later
 - ▶ In that time it had started posting racial slurs, conspiracy theories, and personal attacks
 - ▶ Learned from user interactions (Neff and Nagy 2016)

Outline

Discourse

Text Summarization

Knowledge Graphs

Question Answering

Dialogue Systems

Tools for Legal NLP

Argument Mining

- ▶ Jurafsky and Martin (2020) discuss argument mining; there is not much data or work on it.
- ▶ Peldszus and Stede (2016) provide a (very small) annotated corpus of arguments:
 - ▶ claims
 - ▶ premises
 - ▶ support
 - ▶ attack
 - ▶ types of premises (e.g., from example, from cause to effect, from consequences) (Feng and Hirst 2011)

Scientific Argumentation

Category	Description	Example
AIM	Statement of specific research goal, or hypothesis of current paper	"The aim of this process is to examine the role that training plays in the tagging process"
OWN_METHOD	New Knowledge claim, own work: methods	"In order for it to be useful for our purposes, the following extensions must be made:"
OWN_RESULTS	Measurable/objective outcome of own work	"All the curves have a generally upward trend but always lie far below backoff (51% error rate)"
USE	Other work is used in own work	"We use the framework for the allocation and transfer of control of Whittaker...."
GAP_WEAK	Lack of solution in field, problem with other solutions	"Here, we will produce experimental evidence suggesting that this simple model leads to serious overestimates"
SUPPORT	Other work supports current work or is supported by current work	"Work similar to that described here has been carried out by Merialdo (1994), with broadly similar conclusions."
ANTISUPPORT	Clash with other's results or theory; superiority of own work	"This result challenges the claims of..."

Figure 23.13 Examples for 7 of the 15 labels from the Argumentative Zoning labelset ([Teufel et al., 2009](#)).

- ▶ again, very little labeled data.

Pragmatics

When a diplomat says yes, he means ‘perhaps’;

When he says perhaps, he means ‘no’;

When he says no, he is not a diplomat.

When a diplomat says yes, he means ‘perhaps’;

When he says perhaps, he means ‘no’;

When he says no, he is not a diplomat.

- ▶ language use depends on the context.
 - ▶ e.g. social identity, relationships, setting, conversation history, shared knowledge...
- ▶ this is mostly unexplored in NLP.

Speech Acts

- ▶ Some statements are meant to perform actions (“speech acts”).

“We hold the defendant guilty.”

“I declare them husband and wife.”

“I bequeath this watch to my brother.”

Speech Acts

- ▶ Some statements are meant to perform actions (“speech acts”).

“We hold the defendant guilty.”

“I declare them husband and wife.”

“I bequeath this watch to my brother.”

- ▶ **assertives** = speech acts that commit a speaker to the truth of the expressed proposition, e.g. reciting a creed
- ▶ **directives** = speech acts that are to cause the hearer to take a particular action, e.g. requests, commands and advice
- ▶ **commissives** = speech acts that commit a speaker to some future action, e.g. promises and oaths
- ▶ **expressives** = speech acts that express the speaker’s attitudes and emotions towards the proposition, e.g. congratulations, excuses and thanks
- ▶ **declarations** = speech acts that change the reality in accord with the proposition of the declaration, e.g. baptisms, pronouncing someone guilty or pronouncing someone husband and wife

Speech Acts

- ▶ Some statements are meant to perform actions (“speech acts”).

“We hold the defendant guilty.”

“I declare them husband and wife.”

“I bequeath this watch to my brother.”

- ▶ **assertives** = speech acts that commit a speaker to the truth of the expressed proposition, e.g. reciting a creed
- ▶ **directives** = speech acts that are to cause the hearer to take a particular action, e.g. requests, commands and advice
- ▶ **commissives** = speech acts that commit a speaker to some future action, e.g. promises and oaths
- ▶ **expressives** = speech acts that express the speaker’s attitudes and emotions towards the proposition, e.g. congratulations, excuses and thanks
- ▶ **declarations** = speech acts that change the reality in accord with the proposition of the declaration, e.g. baptisms, pronouncing someone guilty or pronouncing someone husband and wife
- ▶ Important for legal NLP, but hardly any work on this.

"Target Based Speech Act Classification in Political Campaign Text"

Subramanian, Cohn, and Baldwin (2019), N = 258 docs, 6609 sentences:

Utterance	Speech act	Target party	Speaker
Tourism directly and indirectly supports around 38000 jobs in TAS.	<i>assertive</i>	NONE	LABOR
We will invest \$25.4 million to increase forensics and intelligence assets for the Australian Federal Police	<i>commissive-action-specific</i>	LIBERAL	LIBERAL
Labor will prioritise the Metro West project if elected to government.	<i>commissive-action-vague</i>	LABOR	LABOR
A Shorten Labor Government will create 2000 jobs in Adelaide.	<i>commissive-outcome</i>	LABOR	LABOR
Federal Labor today calls on the State Government to commit the final \$75 million to make this project happen.	<i>directive</i>	LIBERAL	LABOR
Good morning everybody.	<i>expressive</i>	NONE	LABOR
The Coalition has already delivered a \$2.5 billion boost to our law enforcement and security agencies.	<i>past-action</i>	LIBERAL	LIBERAL
Malcolm Turnbull's health cuts will rip up to \$1.4 billion out of Australians' pockets every year	<i>verdictive</i>	LIBERAL	LABOR

Table 1: Examples with speech act and target party classes. "Speaker" denotes the party making the utterance.

“Target Based Speech Act Classification in Political Campaign Text”

Subramanian, Cohn, and Baldwin (2019), N = 258 docs, 6609 sentences:

Utterance	Speech act	Target party	Speaker
Tourism directly and indirectly supports around 38000 jobs in TAS.	<i>assertive</i>	NONE	LABOR
We will invest \$25.4 million to increase forensics and intelligence assets for the Australian Federal Police	<i>commissive-action-specific</i>	LIBERAL	LIBERAL
Labor will prioritise the Metro West project if elected to government.	<i>commissive-action-vague</i>	LABOR	LABOR
A Shorten Labor Government will create 2000 jobs in Adelaide.	<i>commissive-outcome</i>	LABOR	LABOR
Federal Labor today calls on the State Government to commit the final \$75 million to make this project happen.	<i>directive</i>	LIBERAL	LABOR
Good morning everybody.	<i>expressive</i>	NONE	LABOR
The Coalition has already delivered a \$2.5 billion boost to our law enforcement and security agencies.	<i>past-action</i>	LIBERAL	LIBERAL
Malcolm Turnbull's health cuts will rip up to \$1.4 billion out of Australians' pockets every year	<i>verdictive</i>	LIBERAL	LABOR

Table 1: Examples with speech act and target party classes. “Speaker” denotes the party making the utterance.

Speech act	%	Kappa (κ)
<i>assertive</i>	40.8	0.85
<i>commissive-action-specific</i>	12.4	0.84
<i>commissive-action-vague</i>	6.6	0.73
<i>commissive-outcome</i>	4.9	0.72
<i>directive</i>	1.7	0.92
<i>expressive</i>	1.9	0.88
<i>past-action</i>	6.3	0.76
<i>verdictive</i>	25.4	0.82

Table 3: Speech act agreement statistics

Speech act	MLP _{ELMo}	Our approach
<i>assertive</i>	0.77	0.80
<i>commissive-action-specific</i>	0.65	0.69
<i>commissive-action-vague</i>	0.45	0.48
<i>commissive-outcome</i>	0.28	0.39
<i>directive</i>	0.58	0.59
<i>expressive</i>	0.55	0.58
<i>past-action</i>	0.45	0.48
<i>verdictive</i>	0.48	0.61

Table 6: Speech act class-wise F1 score.

The World's First Robot Lawyer

The DoNotPay app is the home of the world's first robot lawyer. Fight corporations, beat bureaucracy and sue anyone at the press of a button.

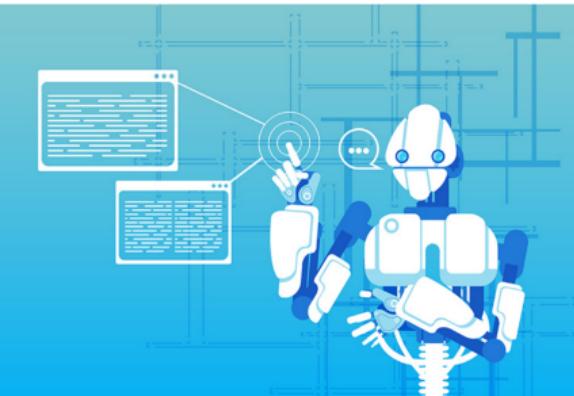
[Sign Up/Login](#)

THINGS YOU CAN DO WITH DONOTPAY

- Fight Corporations
- Beat Bureaucracy
- Find Hidden Money
- Sue Anyone
- Automatically Cancel Your Free Trials



Your Court-Appointed Chatbot – Is Artificial Intelligence Threatening the Legal Profession?



Dangers of Legal NLP Systems

- ▶ (Lack of) Transparency:
 - ▶ Closed-source algorithms result in “black box justice” and could be abused by insiders.
 - ▶ But open-source algorithms are prone to gaming: savvy attorneys could “trick” the algorithm.

Dangers of Legal NLP Systems

- ▶ (Lack of) Transparency:
 - ▶ Closed-source algorithms result in “black box justice” and could be abused by insiders.
 - ▶ But open-source algorithms are prone to gaming: savvy attorneys could “trick” the algorithm.
- ▶ Algorithm can only use evidence that appears in a lot of cases; it ignores special/mitigating circumstances.

Dangers of Legal NLP Systems

- ▶ (Lack of) Transparency:
 - ▶ Closed-source algorithms result in “black box justice” and could be abused by insiders.
 - ▶ But open-source algorithms are prone to gaming: savvy attorneys could “trick” the algorithm.
- ▶ Algorithm can only use evidence that appears in a lot of cases; it ignores special/mitigating circumstances.
- ▶ Current systems would not generalize to new types of cases.
 - ▶ In particular, would not account for new laws/legislation.

Dangers of Legal NLP Systems

- ▶ (Lack of) Transparency:
 - ▶ Closed-source algorithms result in “black box justice” and could be abused by insiders.
 - ▶ But open-source algorithms are prone to gaming: savvy attorneys could “trick” the algorithm.
- ▶ Algorithm can only use evidence that appears in a lot of cases; it ignores special/mitigating circumstances.
- ▶ Current systems would not generalize to new types of cases.
 - ▶ In particular, would not account for new laws/legislation.
- ▶ Teaching a legal NLP system to understand rare evidence, and to understand new laws, would require something much closer to **legal artificial intelligence**.

Legal Vagueness and Value Judgments



- ▶ Even if the AI could read new laws, there is the problem of legal vagueness:
 - ▶ How will the AI decide in this circumstance?

Legal Vagueness and Value Judgments



- ▶ Even if the AI could read new laws, there is the problem of legal vagueness:
 - ▶ How will the AI decide in this circumstance?

- ▶ Making choices in the presence of vagueness or indeterminacy requires value judgements.
 - ▶ What counts as a “good” outcome? Is it even measurable?

