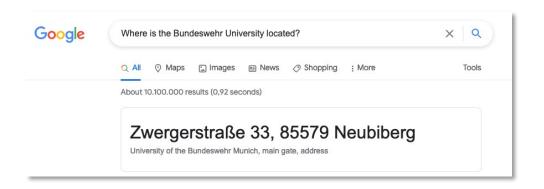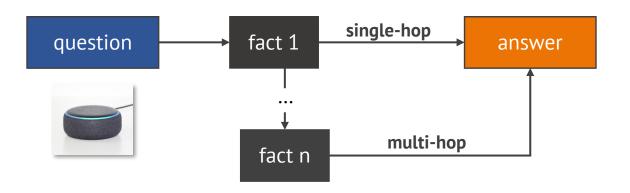Science Track | Early Stage PhD Forum

22nd July 2021

Towards Constructing Multi-hop Reasoning Chains using Local Cohesion

Sabine Ullrich

Michaela Geierhos

Current question answering methods

Single-hop reasoning

- Knowledge-graph approaches

- Simple information extraction

- Answer factual, product-related questions

Multi-hop reasoning

- Build reasoning chains

- At least two steps before conclusion

- Answer more complex questions

▶ Which type of answer is required for which type of question?

Current question answering methods

1. Alexa, Google, Siri: Simple information extraction methods

    Q: Does Kindle support Korean?

    A: Yes, Kindle supports Korean.

2. State-of-the-Art methods: Multi-hop reasoning approaches (eQASC)

    Q: What helps to control **destruction** of **crops**?

    F1: **Insects** **spread disease** and **destroy** **crops**.
    F2: And the industry argues that **harvesting** helps control the **spread** of **insects** and **disease**.

    A: **Harvesting** helps control **destruction** of **crops**.

    ▶ How can we answer more **complex questions**?

Necessity for complex questions answering

3. Complex questions

   Q: How is the situation in Syria?

   A: Critical. ➔ **Not really suffient for this type of question!**

   > A: The situation in Syria is still *critical and unclear*. There is an ongoing civil war since 2011 between the Syrian Arab Republic led by Syrian president Bashar al-Assad and various domestic and foreign forces. Syria's economy is in its worst state since the start of the conflict. (…) More than half of the country's population was displaced. Even as the armed conflict winds down, it is unclear when or if they will be able to return.[1]

   [1]Adapted from https://www.reuters.com/world/middle-east/cost-ten-years-devastating-war-syria-2021-05-26/ and from https://en.wikipedia.org/wiki/Syrian civil war.

▶ How are the sentences in this example connected? What makes the explanation understandable?

Lexical overlapping is not a reliable feature.

> The situation in **Syria** is still *critical and unclear*. **There** is an ongoing **civil war** since 2011 between the Syrian Arab Republic led by Syrian president Bashar al-Assad and various domestic and foreign forces. Syria's economy is in its worst state since the start of **the conflict**. (...) More than **half of the country's population** was displaced. Even as the **armed conflict** winds down, it is unclear when or if **they** will be able to return.

What is a sufficient answer?

- Identify question type: factual vs. open questions

- Answers to factual questions precise and short

- Open questions demand for complex answers

- Longer answers should be cohesive

How can we generate a good answer?

- Identify highly connected sentences

- Form valid reasoning chains

- Find connectives between chain links
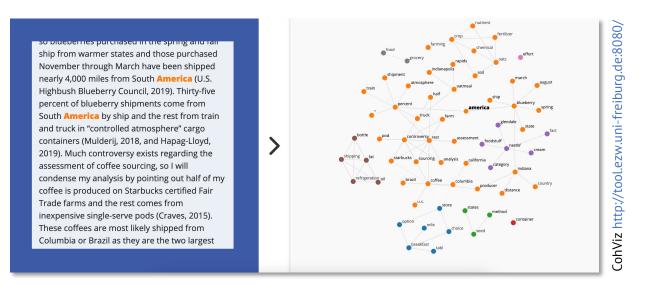
- Coreference resolution for link detection

▶ Most research relies on **overlapping lexical entities.**
Cohesion measures **realistic scenarios** with a **higher lexical and structural variety** across sentences.
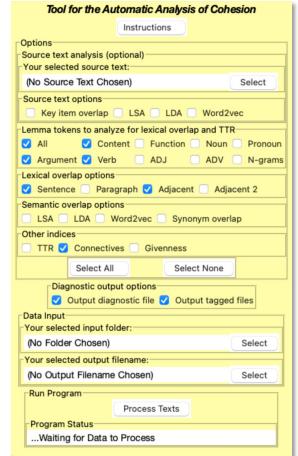
Measuring cohesion can identify connectives.

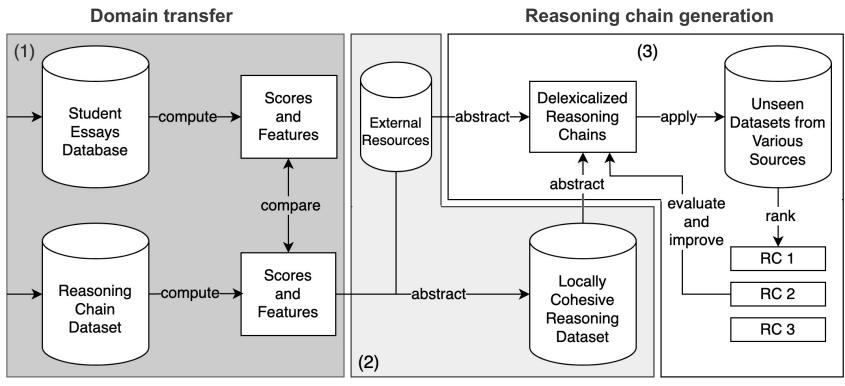**What is cohesion?** → Cohesion measures the connectiveness in a document.

- Global cohesion defines the connectedness on paragraph level

- Local cohesion defines the connectedness on sentence level

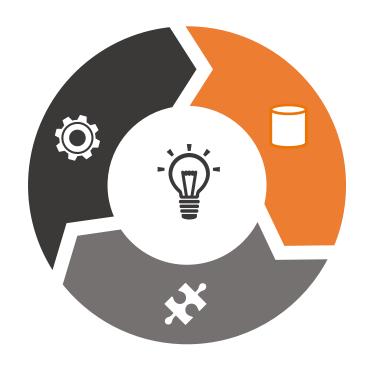▶ The **more cohesive** a text is, the **better understandable** it is for the reader.

...so blueberries purchased in the spring and fall ship from warmer states and those purchased November through March have been shipped nearly 4,000 miles from South **America** (U.S. Highbush Blueberry Council, 2019). Thirty-five percent of blueberry shipments come from South **America** by ship and the rest from train and truck in "controlled atmosphere" cargo containers (Mulderij, 2018, and Hapag-Lloyd, 2019). Much controversy exists regarding the assessment of coffee sourcing, so I will condense my analysis by pointing out half of my coffee is produced on Starbucks certified Fair Trade farms and the rest comes from inexpensive single-serve pods (Craves, 2015). These coffees are most likely shipped from Columbia or Brazil as they are the two largest

CohViz http://tool.ezw.uni-freiburg.de:8080/

TAACO: https://www.linguisticanalysistools.org/taaco.html

Tool for the Automatic Analysis of Cohesion

Our approach – architecture overview

Implementing this architecture, we will have to tackle **3 major challenges**.

Our approach – challenges

1. Domain transfer for cohesive features:

  ▶ Measure cohesion in essays and reasoning chains and determine cohesion for simple sentences.

2. Create a non-overlapping dataset:

  ▶ Create a natural language dataset that represents reasoning chains with high structural and lexical variety.

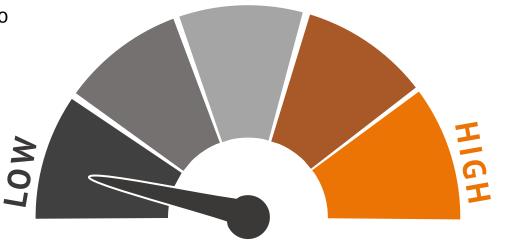**3. Generate delexicalized reasoning chains**:

  ▶ Extract and rank delexicalized reasoning chains from the dataset to represent naturally occurring reasoning chains.

▶ How can cohesion be measured and reasoning chains be extracted from the new dataset?

Domain transfer for cohesive features

Tasks

- Transfer knowledge from AES to reasoning chain evaluation
- Tool for Automatic Analysis of Cohesion (TAACO)
- Compare essays and RCs
- Compare local and global cohesion

Cohesive features

- Connectives
- Type-token-ratio
- Synonym overlap
- (Lexical overlap)
- Pronoun density

LOW

HIGH

MEASURE COHESION

Low cohesion

Reasoning chains are not or little connected. It is difficult to understand the explanation.

High cohesion

Reasoning chains are highly connected. It is easy to follow the explanation.

▶ Use the determined cohesive features to create a dataset.

Creating a non-overlapping dataset

Dataset comparison

- Compare multi-hop datasets for QA
- Compare complexity, length, overlapping…

Lexical adaptation

- Replace words with synonyms
- Possible software: NLTK, WordNet

| Step 01 | Step 02 | Step 03 | Step 04 |

Dataset selection

- Select best suiting dataset
- eQASC, HotpotQA
- Identify overlapping

Structural changes

- Generate paraphrases
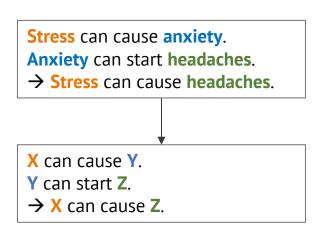- Possible tools and databases: NLPAug, PPDB, PARROT

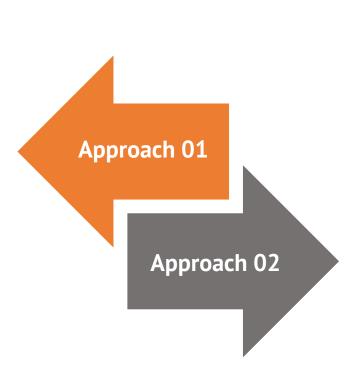▶ Use the dataset to extract delexicalized reasoning chains.
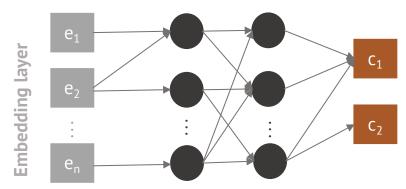
Generating delexicalized reasoning chains

Feed-forward NN

Delexicalization approach

- Replace noun phrases with variables

- POS tagging (NLTK)

- Use predefined set of special tokens

- Graph-based local grammars

Approach 01

Approach 02

Stress can cause anxiety.
Anxiety can start headaches.
→ Stress can cause headaches.

X can cause Y.
Y can start Z.
→ X can cause Z.

Embedding layer

e_1 e_2 e_n

c_1 c_2

Validate candidates

Deep learning approach

- Transformer for encoding (e.g. pre-trained BERT)

- Two-layer feed-forward network with ReLU activation to predict valid reasoning chains
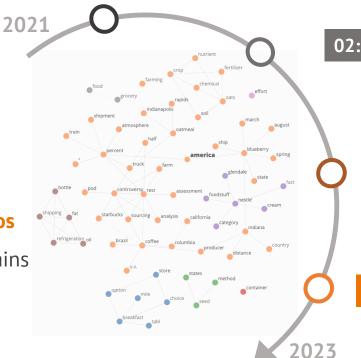
Summary of our research proposal

Multi-hop
Reasoning
Chain
Generation

Our goal is to

- **extract reasoning chains** that

- are taken from **realistic scenarios**

- where connected reasoning chains are **lexically and structurally diverse**.

2021

2023

01: Domain transfer from AES

▶ Calculate cohesion for comprehensible essays and for reasoning chains to analyze a possible domain transfer.

02: Dataset creation

▶ Create a cohesive dataset with highly connected sentences but no lexical overlap.

03: Reasoning chain abstraction

▶ Abstract patterns and delexicalized chains or embeddings from the new dataset.

04: Ranking and evaluation

▶ Rank and evaluate the generalized reasoning chains. The evaluation then helps to further improve the chain extraction.

Thank you for your attention!

Sabine Ullrich & **Michaela Geierhos**
Research Institute CODE
Universität der Bundeswehr München

sabine.ullrich@unibw.de
michaela.geierhos@unibw.de
https://www.unibw.de/datascience