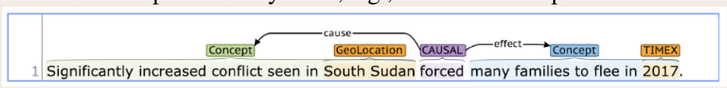
Combining Extraction and Generation for Constructing Belief-Consequence Causal Links

Maria Alexeeva* Allegra A. Beal Cohen† Mihai Surdeanu*

*University of Arizona, † University of Florida

Situating the Task

- Automated event extraction can help computational modelers build models of causes and effects present in systems, e.g., as seen in Sharp et al. 2019



- We aim to enrich models with causal links based on subjective views of populations involved, e.g.:

Political view → Decision to wear a mask

- We frame the task as belief-consequence extraction
- However, in many cases (≈ 49 %), consequences are not explicitly stated:

Belief	Consequence
Political view	Decision to wear a mask
Rice grown in the dry season was thought to have lower risks	Farmers may not need to buy insurance.
Use of chemicals is considered undesirable	X

Approach

Information Extraction:

We extract beliefs (and in future explicit consequences) with a grammar using the Odin information extraction framework (Valenzuela-Escárcega et al. 2015):

```
- name: belief-rule
  label: Belief
  type: dependency
  pattern: |
    trigger = [lemma=/consider/]
    believer:Agent = /nsubj/
    belief:Proposition = /xcomp/
```

Text Generation:

We generate implicit consequences using the OpenAI InstructGPT-3 model (Ouyang et al. 2022).

We prompt the model with several few-shot examples, e.g.:

Belief: *Rice grown in the dry season produced higher yields and was perceived to have lower risks.*

Consequence: *Farmers may not need to buy insurance for rice grown during the dry season.*

Text Generation and Human Annotators

For each of the 18 extracted beliefs that did not have explicit consequences within a one-paragraph context window, we generate three potential consequences. We judge generation successful if at least one out of three automatic predictions matches the human-generated consequences. We observe the following overlap in content:

Condition	Overlap
two annotators	13 (72%)
GPT-3 and one annotator	12 (66%)
GPT-3 and both annotators	9 (50%)

Belief: *Planners and technicians feel that the development of irrigation systems could offer a solution to the crisis in food production in Africa.*

Annotator 1: *Planners and officials will invest more in the development of irrigation systems.*

Annotator 2: *They should develop irrigation systems.*

GPT-3: *Planners and technicians focus on the development of irrigation systems.*

Evaluation of Text Generation Independent of Human Annotators

Additionally, we evaluate generations in terms of their appropriateness for the prompts. Some of the error types observed in belief consequences generated by GPT-3:

	Belief	Consequence	Error Type
1.	women unanimously believe their work to be profitable and well worth pursuing.	Women continue to engage in embroidery work.	too specific
2.	Mothers also believe that a child does not need to eat solid food before 8 months.	Breastfeeding rates are low in the study area.	contradiction
3.	farmers strongly feel that increased bird pressure intensity with resulting higher bird damages are the main reason for these lower yields in 2006.	Farmers stop growing rice.	wrong focus
4.	gifts, financial support, and physical attention are themselves considered important demonstrations of romantic love.	Men in this culture often give gifts, offer financial support, and provide physical attention to more than one woman at a time.	unsubstantiated consequence
5.	personnel believe that up to an additional 60,000 ha could be rehabilitated.	The additional 60,000 ha will not be rehabilitated.	wrong polarity
6.	use of chemicals and machinery on their paddy field is often considered	Use of chemicals and machinery on their paddy field is often considered	restatement

Zero-shot

While the model generated chunks of text that could be viewed as consequences even without explicit examples, in those cases it also continued to. . .

- Generate text to reach the word limit:

Belief: people believe that their hard work in the rice paddies is part of a covenant with their supreme deity. That's why

Consequence: they work so hard." "I see. So their hard work is a form of worship." "Exactly

- Generate text that was too specific:

Belief: [...] centers are often considered the best private family planning clinics. That's why

Consequence: the Trump administration's proposed rule to cut $300 million in Title X family planning funding is so alarming. Title X

Generation consistency

The quality of several consequences generated for each belief is not necessarily consistent. Below are the counts of beliefs for which all three generated consequences were correct, some were correct, or none were correct:

Condition	Count
all correct	8
a mix of correct and incorrect	7
all incorrect	3

Future Work

- Extracting explicit consequences
- Fine-tuning language models to increase consistency of generated consequences
- Collecting data
- Using the extracted belief-consequence pairs to build cognitive models

Acknowledgements

The authors thank the anonymous reviewers for helpful discussion and Andrew Zupon for help with annotation.

This work was supported by the Defense Advanced Research Projects Agency (DARPA) under the Habitus program. Maria Alexeeva and Mihai Surdeanu declare a financial interest in lum.ai. This interest has been properly disclosed to the University of Arizona Institutional Review Committee and is managed in accordance with its conflict of interest policies.