

Multilingual Generative Language Models for Zero-Shot Cross-Lingual Event Argument Extraction

Kuan-Hao Huang^{*1}, I-Hung Hsu^{*2}, Premkumar Natarajan², Kai-Wei Chang¹, Nanyun Peng^{1,2}

¹ University of California, Los Angeles

² Information Science Institute, University of Southern California



Event Argument Extraction

- Goal: extract the participants for an event described in a sentence
- Input: sentence, event trigger
- Output: (role, argument) pairs
- Challenge:
 - Dependency
 - Event structure

Agent	coalition
Victim	civilians, woman
Instrument	missile
Place	houses

Five Iraqi civilians, including a woman, were killed Monday when their houses were hit by a missile fired by the US - led coalition warplanes, witnesses said.

Trigger for a Life:Die event

Zero-Shot Cross-Lingual Event Argument Extraction

- Training examples come from source languages
- Testing examples come from target languages
- Challenge:
 - Discrepancy between language properties
 - Different vocabularies, different word order and grammar, etc.

Agent	coalition
Victim	civilians, woman
Instrument	missile
Place	houses

Agent	以軍
Victim	青年
Instrument	催泪弹, 子弹, 实弹
Place	None

Five Iraqi civilians, including a woman, were killed Monday when their houses were hit by a missile fired by the US - led coalition warplanes, witnesses said.

巴勒斯坦人持续以石块攻击以色列的部队，以军则是还以催泪弹、橡皮子弹甚至是实弹，结果又造成两名巴勒斯坦青年丧生，10多人受伤。

Previous Approaches

- Most previous approaches are classification-based models
 - GATE [Ahmad+ 2021] and CL-GCN [Subburathinam+ 2019]
- Recently, several works have shown that generation-based models perform better than classification-based models
 - DEGREE [Hsu+ 2022], BART-Gen [Li+ 2021], and TANL [Paolini+ 2021]
 - Strong performance, but specific to the monolingual setting

Agent	coalition
Victim	civilians, woman
Instrument	missile
Place	houses

Language-dependent!

Autoregressive generation considers dependency

The coalition led civilians and woman died by using missile in their houses.

Prompts provide event structure

Five Iraqi civilians, including a woman, were killed Monday when their houses were hit by a missile fired by the US - led coalition warplanes, witnesses said.

Trigger is killed. Somebody led some victim died by using some way in somewhere.

Our Goal

- We aim to explore the possibility of applying generation-based models for zero-shot cross-lingual event argument extraction
 - Output design
 - Prompt design

Proposed Method: X-Gear - Language-Agnostic Output Format

Easy to decode the final predictions

Language-agnostic format for language discrepancy

<Agent> coalition </Agent> <Victim> civilians [and] woman </Victim>
<Instrument> missile </Instrument> <Place> houses </Place>

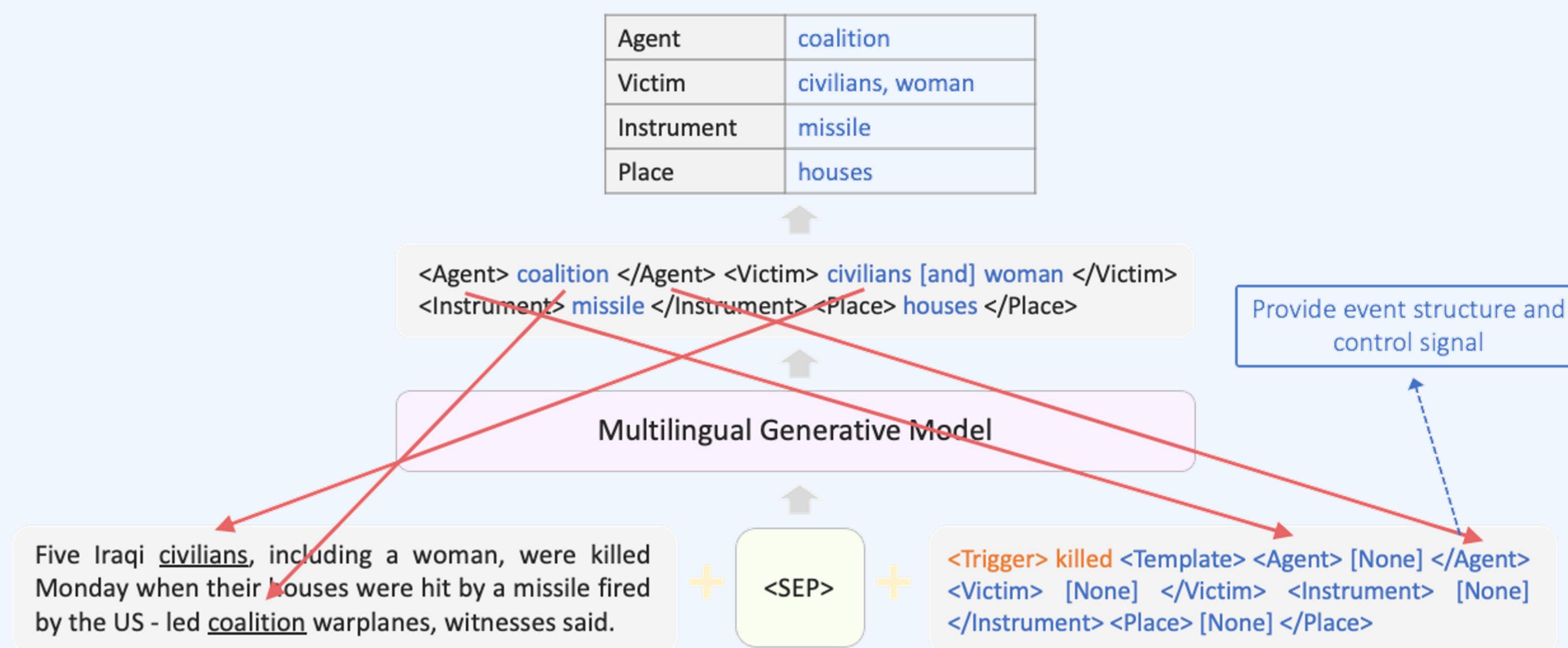
Multilingual Generative Model

Five Iraqi civilians, including a woman, were killed Monday when their houses were hit by a missile fired by the US - led coalition warplanes, witnesses said.

<SEP> + Prompt

They are special tokens and don't belong to any language

Proposed Method: X-Gear - Prompt Design and Copy Mechanism



Main Experiments

Argument Classification F1 on ACE-2005

Model	# of parameters	en ↓ xx	en ↓ xx	en ↓ ar	ar ↓ en	ar ↓ ar	zh ↓ zh	zh ↓ zh	zh ↓ ar	avg	
OneIE (XLM-R-large) (Lin et al., 2020)	~570M	63.6	42.5	37.5	57.8	27.5	31.2	69.6	51.5	31.1	45.8
CL-GCN (XLM-R-large) (Subburathinam et al., 2019)	~570M	59.8	29.4	25.0	47.5	25.4	19.4	62.2	40.8	23.3	37.0
GATE (XLM-R-large) (Ahmad et al., 2021)	~590M	67.0	49.2	44.5	59.6	27.6	26.3	70.6	46.7	37.3	47.6
GATE (mBART-50-large)	~630M	65.5	43.0	38.9	58.5	27.5	26.1	65.9	45.3	30.2	44.5
GATE (mT5-base)	~590M	59.8	47.7	32.6	45.4	20.7	21.0	64.0	35.3	22.8	38.8
TANL (mT5-base) (Paolini et al., 2021)	~580M	59.1	38.6	29.7	50.1	18.3	16.9	65.2	33.3	18.3	36.6
X-GEAR (mBART-50-large)	~610M	68.3	48.9	37.8	59.8	30.5	29.2	63.6	45.9	32.3	46.2
X-GEAR (mT5-base)	~580M	67.9	53.1	42.0	66.2	27.6	30.5	69.4	52.8	32.0	49.1
X-GEAR (mT5-large)	~1230M	71.2	54.0	44.8	68.9	32.1	33.3	68.9	55.8	33.1	51.3

Ablation Study on Copy Mechanism

Model	en ↓ xx	ar ↓ xx	zh ↓ xx	xx ↓ en	xx ↓ ar	xx ↓ zh	avg
mBART-50-large	51.6	39.8	47.2	48.2	43.2	47.2	46.2
- w/o copy	50.9	42.2	49.6	50.6	43.5	48.7	47.6
mT5-base	54.3	41.4	51.4	49.4	46.7	51.0	49.1
- w/o copy	52.1	39.5	47.6	48.1	42.7	48.5	46.4
mT5-large	56.7	44.8	52.6	53.0	48.9	52.1	51.3
- w/o copy	55.1	45.0	51.5	52.0	46.3	53.2	50.5

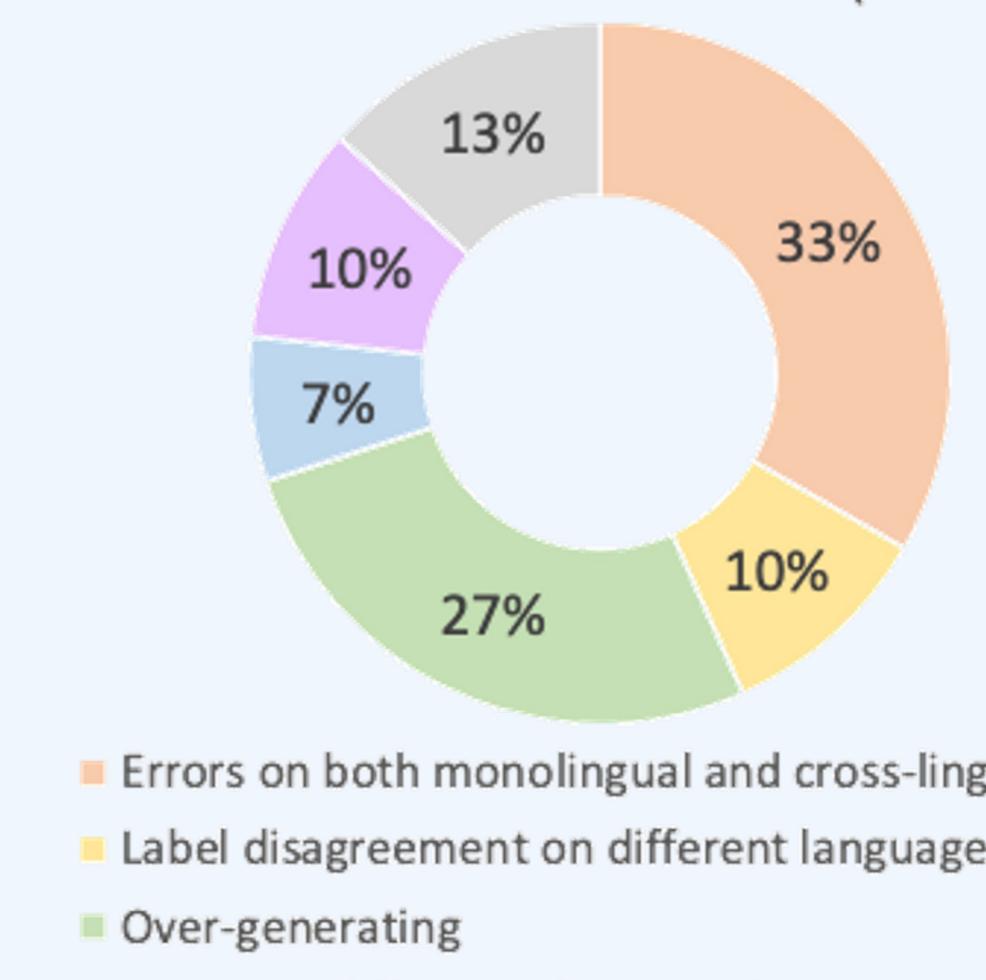
Including event type in prompts

Five Iraqi civilians, including a woman, were killed Monday when their houses were hit by a missile fired by the US - led coalition warplanes, witnesses said.

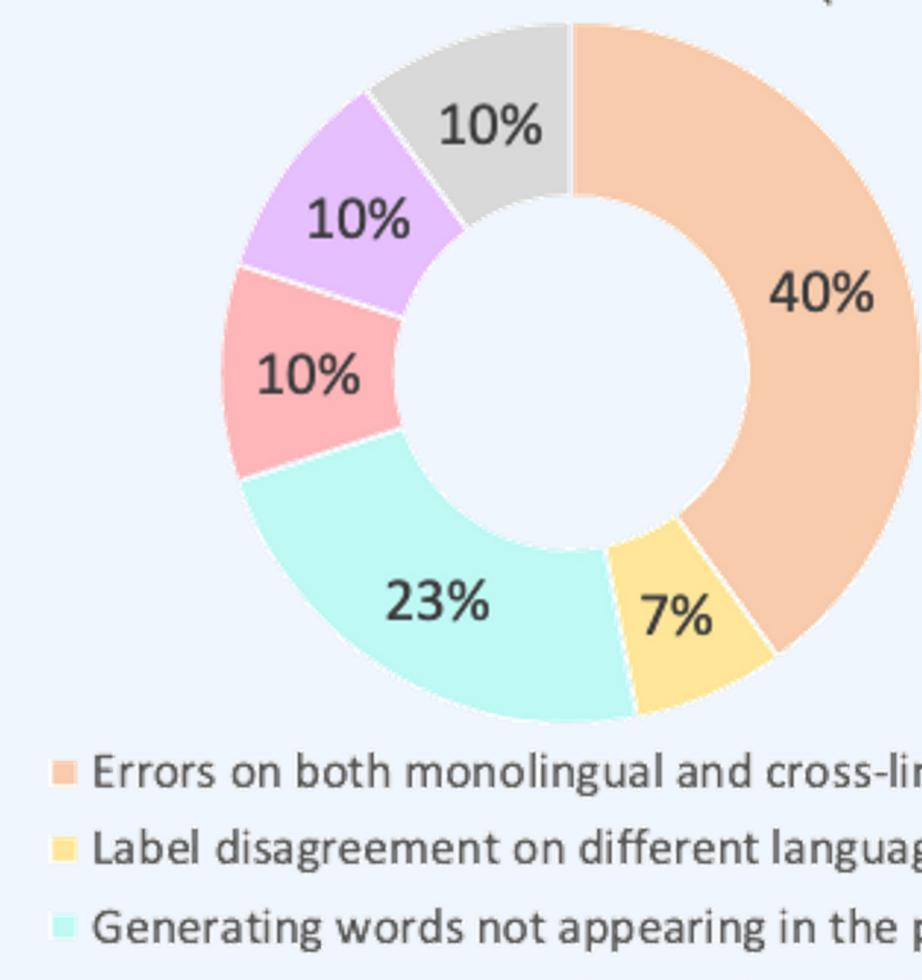
Model	en ↓ xx	ar ↓ xx	zh ↓ xx	xx ↓ en	xx ↓ ar	xx ↓ zh	avg
X-GEAR (mT5-base)	54.3	41.4	51.4	49.4	46.7	51.0	49.1
w/ English Tokens	53.3	39.3	52.3	49.2	46.5	49.2	48.3
w/ Translated Tokens	51.7	40.4	52.2	49.8	45.6	48.8	48.1
w/ Special Tokens	52.3	39.7	51.8	49.0	45.4	49.3	47.9

Error Analysis

Error Distribution for X-GEAR (ar ⇒ en)



Error Distribution for X-GEAR (zh ⇒ en)



Conclusion

- We propose X-Gear, a generation-based model for zero-shot cross-lingual event argument extraction
 - Inherit the benefits of generation-based models
 - Language-agnostic templates
 - Copy mechanism
- Significant improvements over previous baselines

Code is available at

<https://github.com/PlusLabNLP/X-Gear>

<https://github.com/PlusLabNLP/TextEE>

