# **ATOMIC**: An **AT**las **O**f **M**ach**I**ne **C**ommonsense for If-Then Reasoning

KV Yaswanth
VQ63094

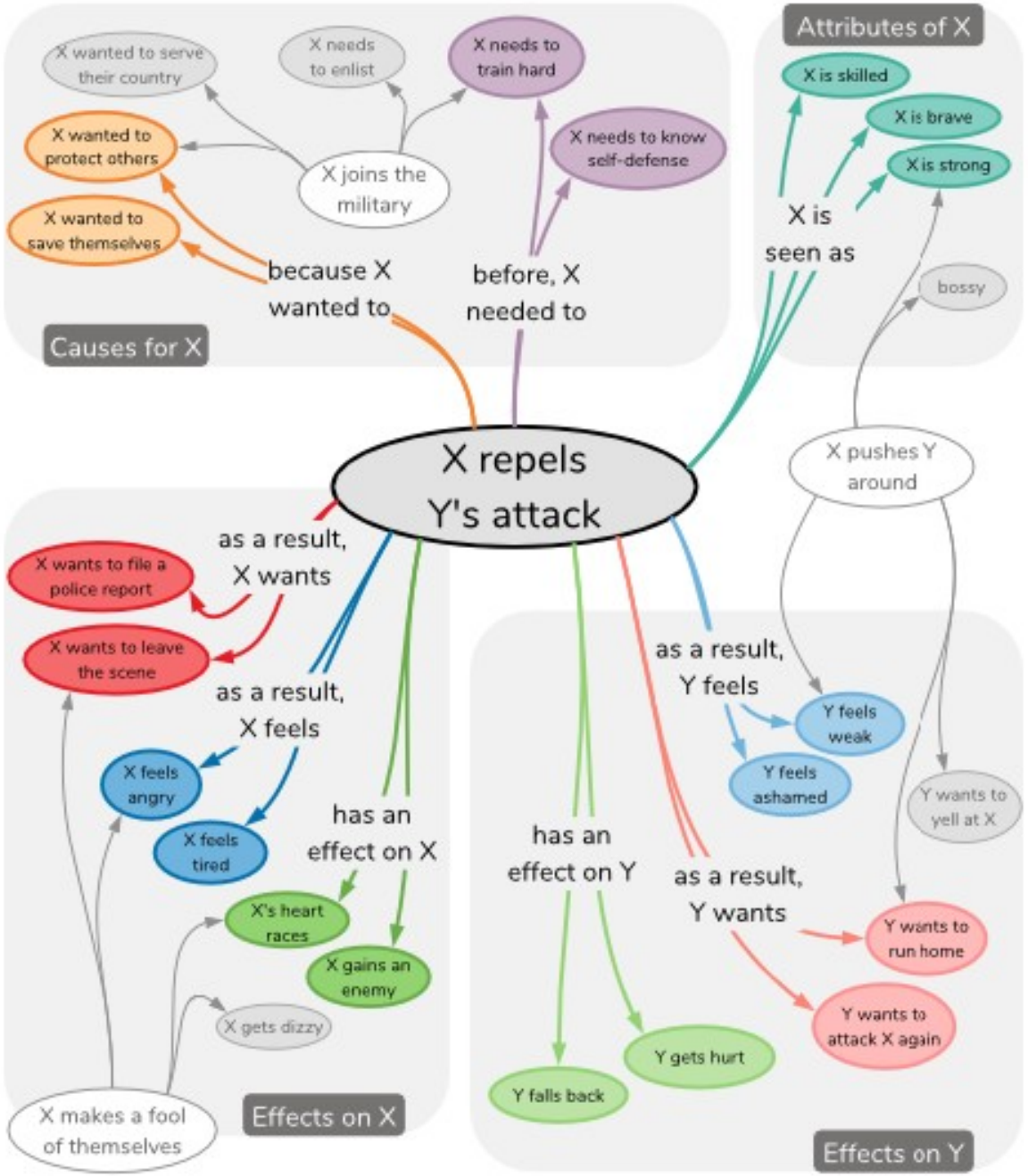
# Introduction

- ATOMIC, a large-scale knowledge graph of commonsense reasoning
- inferential knowledge through typed if-then relations
- Addresses the gap in machine commonsense for AI systems
- Aids in understanding causes, effects, and mental states related to everyday events

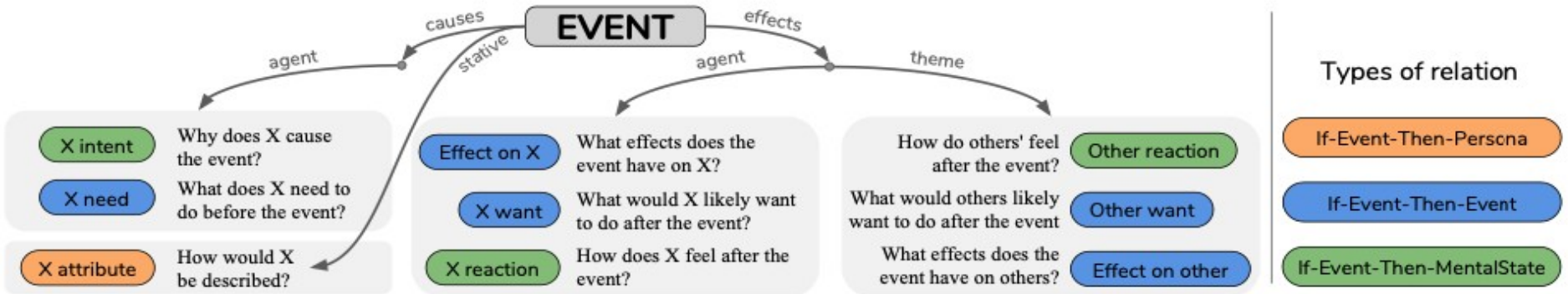
# High-Level Summary

- 877K instances of inferential knowledge from 300K events
- Nine if-then reasoning types, including intents, reactions, needs, and attributes
- Taxonomy of relationships (If-Event-Then-Mental-State, If-Event-Then-Event, If-Event-Then-Persona)
- Multitask learning approaches for inference generation

# Consider an event



# Reasoning Types



# Nine Inferential Dimensions

| Event | Type of relations | Inference examples | Inference dim. |
|---|---|---|---|
| "PersonX pays PersonY a compliment" | If-Event-Then-Mental-State | PersonX wanted to be nice<br>PersonX will feel good<br>PersonY will feel flattered | xIntent<br>xReact<br>oReact |
| | If-Event-Then-Event | PersonX will want to chat with PersonY<br>PersonY will smile<br>PersonY will compliment PersonX back | xWant<br>oEffect<br>oWant |
| | If-Event-Then-Persona | PersonX is flattering<br>PersonX is caring | xAttr<br>xAttr |
| "PersonX makes PersonY's coffee" | If-Event-Then-Mental-State | PersonX wanted to be helpful<br>PersonY will be appreciative<br>PersonY will be grateful | xIntent<br>oReact<br>oReact |
| | If-Event-Then-Event | PersonX needs to put the coffee in the filter<br>PersonX gets thanked<br>PersonX adds cream and sugar | xNeed<br>xEffect<br>xWant |
| | If-Event-Then-Persona | PersonX is helpful<br>PersonX is deferential | xAttr<br>xAttr |
| "PersonX calls the police" | If-Event-Then-Mental-State | PersonX wants to report a crime<br>Others feel worried | xIntent<br>oReact |
| | If-Event-Then-Event | PersonX needs to dial 911<br>PersonX wants to explain everything to the police<br>PersonX starts to panic<br>Others want to dispatch some officers | xNeed<br>xWant<br>xEffect<br>oWant |
| | If-Event-Then-Persona | PersonX is lawful<br>PersonX is responsible | xAttr<br>xAttr |

# Free Form Crowdsourced Data

# Single vs. Multitask Learning

- 9ENC9DEC
- EVENT2(IN)VOLUNTARY
- EVENT2PERSONX/Y
- EVENT2PRE/POST

# BLEU Scores

| Dataset | Model | xIntent | xNeed | xAttr | xEffect | xReact | xWant | oEffect | oReact | oWant |
|---|---|---|---|---|---|---|---|---|---|---|
| DEV | 9ENC9DEC | **8.35** | 17.68 | 5.18 | **10.64** | **5.38** | **13.24** | 6.49 | 5.17 | 12.08 |
| | NearestNeighbor | 6.14 | 11.36 | 3.57 | 5.81 | 4.37 | 7.73 | **8.02** | **6.38** | 8.94 |
| | EVENT2(IN)VOLUNTARY | 7.51 | **17.80** | 5.18 | 10.51 | 4.78 | 12.76 | 7.04 | 4.84 | **12.48** |
| | EVENT2PERSONX/Y | 7.31 | 17.08 | **5.26** | 9.78 | 4.83 | 12.14 | 6.38 | 4.84 | 11.45 |
| | EVENT2PRE/POST | 7.58 | 17.17 | – | 10.50 | 4.73 | 11.78 | 6.71 | 4.87 | 11.52 |
| TEST | 9ENC9DEC | **8.68** | 18.15 | **5.18** | **10.34** | **5.43** | **14.50** | 6.61 | 5.08 | 12.73 |
| | NearestNeighbor | 6.64 | 11.35 | 3.37 | 5.52 | 4.59 | 8.17 | **7.58** | **5.88** | 9.18 |
| | EVENT2(IN)VOLUNTARY | 7.94 | **18.22** | 5.02 | 9.78 | 4.78 | 13.67 | 7.16 | 4.71 | **13.23** |
| | EVENT2PERSONX/Y | 7.67 | 17.33 | 5.09 | 9.45 | 4.82 | 13.19 | 6.59 | 4.68 | 11.70 |
| | EVENT2PRE/POST | 7.96 | 17.42 | – | 9.79 | 4.75 | 12.85 | 6.90 | 4.76 | 11.97 |

# Human Evaluation

| Model | xNeed | xIntent | xAttr | xEffect | xReact | xWant | oEffect | oReact | oWant | average |
|---|---|---|---|---|---|---|---|---|---|---|
| 9ENC9DEC | 48.74 | 51.70 | 52.20 | 47.52 | 63.57 | 51.56 | 22.92 | 32.92 | 35.50 | 45.32 |
| EVENT2(IN)VOLUNTARY | 49.82 | **61.32** | 52.58 | 46.76 | 71.22 | 52.44 | **26.46** | **36.04** | 34.70 | **47.93** |
| EVENT2PERSONX/Y | **54.04** | 53.93 | **52.98** | **48.86** | 66.42 | **54.04** | 24.72 | 33.80 | 35.08 | 46.41 |
| EVENT2PRE/POST | 47.94 | 57.77 | 52.20 | 46.78 | **72.22** | 47.94 | 26.26 | 34.48 | 35.78 | 46.76 |
| gold ATOMIC annotations | 81.98 | 91.37 | 78.44 | 83.92 | 95.18 | 90.90 | 84.62 | 86.13 | 83.12 | 86.18 |

# Strengths

- Focuses on inferential knowledge, unlike taxonomic datasets
- Largest dataset of its kind, with validated human annotations.
- Enables AI systems to predict plausible causes, effects, and motivations for unseen events.

# Weaknesses

- Crowdsourced data might include biases or noise, despite validation.
- Focuses mainly on single-step inferences and lacks multi-turn reasoning for complex narratives.
- While innovative, there is minimal overlap with datasets like ConceptNet (~7%), raising integration questions.

# Story Generation and Interactive Fiction

- **Story Generation**:
  - Use ATOMIC to model characters' motivations, actions, and consequences dynamically.
  - Enhance plot consistency through plausible event sequences.
- **Interactive Fiction**:
  - Implement character-driven decision-making (e.g., emotional reactions or needs).
  - Enable AI to adapt storylines based on player actions.
- **Future Use Cases**:
  - Dynamic branching narratives where NPCs' actions/reactions feel realistic.
  - Building richer character backstories by inferring unstated motivations.

# Example Application

- **Scenario**:
  - Player chooses to “save an NPC.”
  - Using ATOMIC:
    - Predict NPC’s gratitude or next steps (e.g., offering a reward).
    - Generate subsequent events influenced by inferred motives (e.g., NPC accompanies the player).
- **Interactive Fiction Enhancement**:
  - Enables more immersive storytelling by modeling plausible cause-effect relationships.

thank you