Commonsense Properties from Query Logs and Question Answering Forums

Julien Romero, Simon Razniewski, Koninika Pal, Jeff Z. Pan, Archit Sakhadeo, Gerhard Weikum
Goal

- **Mine Commonsense Knowledge (CSK) about:**
  - Object properties
  - Human behavior
  - General concepts

- **Focus on salient properties like**
  - (bananas, are, edible)
  - (children, like, bananas)

- **Avoid non salient properties like (from ConceptNet)**
  - (elephant, CapableOf, visit the grocery store)
  - (dog, HasProperty, one among many animals)
Applications

- Chatbot
  - Me: Hi Pandora, what do you suggest for breakfast?
  - Her: What about bouillabaisse for a starter?

- (Visual) Question Answering
  - Q: What’s taller, the giraffe or the mountain?
  - A: The giraffe

- Visual content understanding

- Queries Interpretation
  - Jordan weather next week
Challenges

■ Rarely expressed
■ Non-encyclopedic (no Wikipedia)

Banana

From Wikipedia, the free encyclopedia

A banana is an edible fruit – botanically a berry – produced by several kinds of large herbaceous flowering plants in the genus *Musa*. In some countries, bananas used for cooking may be called “plantains”, distinguishing them from dessert bananas. The fruit is variable in size, color, and firmness, but is usually elongated and curved, with soft flesh rich in starch covered with a rind, which may be green, yellow, red, purple, or brown when ripe. The fruits grow in clusters hanging from the top of the plant. Almost all modern edible seedless (parthenocarp) bananas come from two wild species – *Musa acuminata* and *Musa balbisiana*. The scientific names of most cultivated bananas are *Musa acuminata*, *Musa balbisiana*, and *Musa × paradisiaca* for the hybrid *Musa acuminata × M. balbisiana*, depending on their genomic constitution. The old scientific name for this hybrid, *Musa sapientum*, is no longer used.

■ Noise and high bias on online content
Previous Work

- Traditional Knowledge Bases (Wikidata, DBpedia, Yago, ...)
  - No commonsense
- ConceptNet, BT Technology Journal, 2004
  - Manual, does not scale
- Webchild, MPI, WSDM 2014
  - Focus on possible properties, not salient ones
- TupleKB, AllenAI, TACL 2017
  - Domain specific
General Pipeline

- Query Logs
- Candidate Gathering
- Corroboration
- Ranking
- Grouping

Sources:
- QA Forums
- Encyclopedias
- Books
- Answer Snippets
- Image Tags

Result:
QUASIMODO Common-sense KB
Candidate Gathering

- Main idea: Extract facts from questions
  - When asking a question, make assumptions about the world

Why are bananas yellow? Bananas are yellow!

- Harvest human curiosity, « wisdom of the crowds »
Candidate Gathering – Query Logs

- Indirect access to the query logs through autocompletion

```
why do cats
why do cats purr
why do cats like boxes
why do cats meow
why do cats knead
why do cats sleep so much
why do cats hate water
why do cats like catnip
why do cats lick you
why do cats have whiskers
```
Candidate Gathering – QA Forums

Yahoo! Answers (research datasets) → why-how questions → Answers.com (sitemap)

why-how questions → Quora (semi-manually) → Reddit (dump)
### Candidate Gathering – Statistics

| Pattern       | In Query Logs | In QA Forums |
|---------------|---------------|--------------|
| how does      | 19.4%         | 7.5%         |
| why is        | 15.8%         | 10.4%        |
| how do        | 14.9%         | 38.07%       |
| why do        | 10.6%         | 9.21%        |
| how is        | 10.1%         | 4.31%        |
| why does      | 8.97%         | 5.46%        |
| why are       | 8.68%         | 5.12%        |
| how are       | 5.51%         | 1.8%         |
| how can       | 3.53%         | 10.95%       |
| why can’t     | 1.77%         | 1.40%        |
| why can       | 0.81%         | 0.36%        |
Candidate Gathering – Results

Questions to statements to tuples using OpenIE

Why are lions hunting zebras?

Q2S

Lions are hunting zebras

OpenIE

(lions, are hunting, zebras)

Normalize

(lion, hunt, zebras)
Corroboration

- Reduce noise thanks to additional signals from:
  - Wikipedia and Simple Wikipedia
  - Answer snippets from search engines
  - Google Books
  - Image Tags from OpenImages and Flickr
  - Google’s Conceptual Captions dataset

Wildlife Photographer of the Year award goes to Yongqing Bao for image of Tibetan fox attacking marmot

Train classifier from all signals in 700 manually annotated triples
Ranking

- From Corroboration, get plausibility score $\pi$
- Define probability:
  \[
P[s, p, o] = \frac{\pi(spo)}{\sum_{x \in KB} \pi(x)}
  \]
- Derive typicality $\tau$ and saliency $\sigma$:
  \[
  \tau(s, p, o) = P[p, o \mid s] = \frac{P[s, p, o]}{P[s]}
  \]
  \[
  \sigma(s, p, o) = P[s \mid p, o] = \frac{P[s, p, o]}{P[p, o]}
  \]
Grouping

- Reduce redundancy
- Co-clustering method based on tri-factorization
- Compute clusters for SO pairs and clusters for P phrases and align them with each other when meaningful
- Number of (soft) clusters for SO pairs and for P phrases can be different

| P cluster                          | SO cluster                        |
|-----------------------------------|-----------------------------------|
| make noise at, be loud at, croak in | fox-night, frog-night, donkey-night |
| sleep in, be bored in, talk in    | student-class, student-lectures   |
## Statistics

|                     | Full KB |          |          |          |
|---------------------|---------|----------|----------|----------|
|                     | #S      | #P       | #P ≥ 10  | #SPO    | #SPO/S  |
| ConceptNet-full@en  | 842,532 | 39       | 39       | 1,334,425 | 1.6     |
| ConceptNet-CSK@en   | 41,331  | 19       | 19       | 214,606  | 5.2     |
| TupleKB             | 28,078  | 1,605    | 1,009    | 282,594  | 10.1    |
| WebChild            | 55,036  | 20       | 20       | 13,323,132 | 242.1  |
| Quasimodo           | 80,145  | 78,636   | 6084     | 2,262,109 | 28.2    |

|                     | animals | occupations |
|---------------------|---------|-------------|
|                     | #S      | #SPO | #S | #SPO |
| ConceptNet-full@en  | 50      | 2,678 | 50 | 1,906 |
| ConceptNet-CSK@en   | 50      | 1,841 | 50 | 1,495 |
| TupleKB             | 49      | 16,052 | 38 | 5,321 |
| WebChild            | 50      | 27,223 | 50 | 26,257 |
| Quasimodo           | 50      | 39,710 | 50 | 18,212 |
## Anecdotal Examples

| Category                                      | Example                  |
|-----------------------------------------------|--------------------------|
| Practical knowledge from human                | (car, slip on, ice)      |
| Problems linked to a subject                  | (pen, can, leak)         |
| Emotions linked to events                     | (divorce, can, hurt)     |
| Human behaviors                               | (ghost, scare, people)   |
| Visual facts                                  | (road, has_color, black) |
| Cultural knowledge (here U.S.)                | (school, have, locker)   |
| Comparative knowledge                         | (light, faster than, sound) |
Precision – Same Subjects

Sample from a list of common subjects (most popular animals and occupations)

![Graph showing quality@5 for meaningfulness, typicality, and saliency for different systems: ConceptNet, WebChild, TupleKB, Quasimodo.](image-url)
Recall

- Given a subject, ask MTurks to give a statement begin by “Subjects …”, like “Elephants are grey”
- Strict = exact match, Relaxed = partial match
**Multiple Choice Question Answering**

Where would I not want a fox?
- hen house
- England
- mountains
- English hunt
- California

|          | KB       | All       |
|----------|----------|-----------|
| Questions (Train/Test) | 10974/3659 |
| Random   | 22.0     |
| word2vec | 27.2     |
| Quasimodo| 31.3     |
| ConceptNet | 27.5    |
| TupleKB  | 27.5     |
| WebChild | 24.1     |
Conclusion

■ New methods for acquiring CSK from non-standard sources
■ Better coverage of salient properties
■ Extrinsic evaluations for multiple choice QA
■ Data and code available: github.com/Aunsiels/CSK
Additional slides
Precision – Entire CSKs

Sample from the entire knowledge base, on all subjects

| quality@5 | meaningfulness | typicality | saliency |
|-----------|----------------|------------|----------|
| 4.5       |                |            |          |
| 4.0       |                |            |          |
| 3.5       |                |            |          |
| 3.0       |                |            |          |
| 2.5       |                |            |          |

ConceptNet | WebChild | TupleKB | Q’modo-τ | Q’modo-σ
Future Work

- Cultural knowledge
- Study of stereotypes
- Temporal evolution of the knowledge base
- Improve ranking methods
- Scale to the entire web
Useful Links

- Data: https://www.mpi-inf.mpg.de/departments/databases-and-information-systems/research/yago-naga/commonsense/quasimodo/
- Code: https://github.com/Aunsiels/CSK
- http://conceptnet.io/
- http://data.allenai.org/tuple-kb/
- https://www.mpi-inf.mpg.de/departments/databases-and-information-systems/research/yago-naga/commonsense/webchild/