A Survey of Crowdsourcing Methods for Commonsense Knowledge Collection

Author: Ilinca Elena Ioana Rentea (I.E.I.Rentea-1@student.tudelft.nl)
Supervisors: Gaole He, Ujwal Gadiraju, Jie Yang

TU Delft

1. Background

- Common sense knowledge = set of information typically possessed by all humans; e.g. "if you stay outside while raining, you will get wet"
- Typically excluded from oral/written communication [2]
- Necessary to minimize gap between machine learning models and humans performance
- Key method of knowledge collection is crowdsourcing, directly from source
- No existing work systematically surveys and compares crowdsourcing methods for commonsense knowledge collection

2. Research Question

- What do existing crowdsourcing methods do to collect commonsense knowledge?
- How efficient, costly ,and accurate are existing crowdsourcing methods to collect commonsense knowledge?

3. Method and process

- Systematic literature survey on crowdsourcing methods for common sense knowledge collection
- Compare methods based on:
  - Efficiency - throughput
  - Cost - average cost per knowledge tuple
  - Quality- average accuracy per knowledge tuple

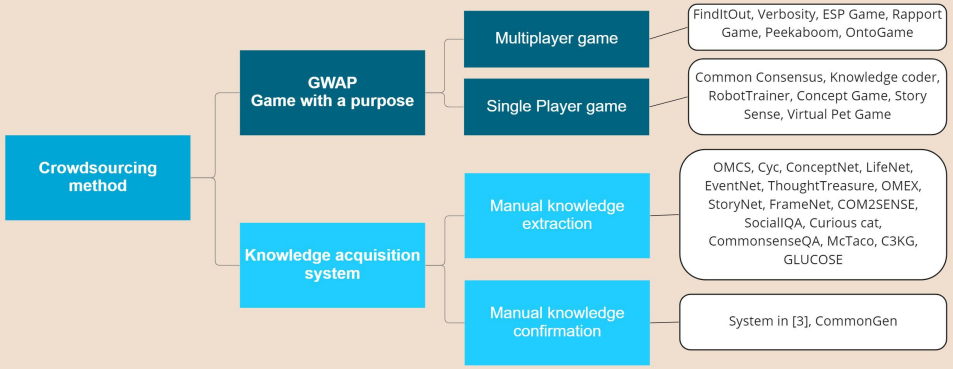

Figure 1: Taxonomy of crowdsourcing methods for common sense knowledge collection

4. Results

- Taxonomy from Figure 1 was created based on the surveyed methods which were compared within each sub-category based on different criteria such as type of knowledge collected, purpose of system ,or intended user
- Methods from Figure 1 were compared based on efficiency, cost ,and quality of collected knowledge using Table 1 which contains values gathered from reports.

5. Conclusion & Future Work

- Collected values (see table 1) suggest that most games with a purpose are as efficient ,and collect data as accurate as knowledge acquisition systems, while having lower costs.
- Future work: include more crowdsourcing systems and collect analysis data for all methods

Method	Throughput	Cost	Accuracy
Robot Trainer	12	-	63.14%
Common Consensus	38880	-	-
Knowledge Coder	13	-	-
Concept Game	500	-	-
Virtual Pet Game	2796	-	92.07%
Story Sense	124	-	86.55%
FindItOut	20016	-	95.6%
Verbosity	1124	-	85%
The ESP Game	10337	-	85%
Rapport Game	76	-	-
Peekaboom	36225	-	100%
OntoGame	-	-	99%
OMCS-1	651	-	75%
OMCS-2	-	-	85%
Cyc	-	-	-
ConceptNet	-	-	68%
LifeNet	-	-	89%
EventNet	-	-	62%
ThoughtTreasure	28	-	-
OMEX	11520		62%
StoryNet	-	-	-
FrameNet	-	-	-
COM2SENSE	-	-	95%
SocialIQA	-	-	87%
Curious Cat	21	-	96%
CommonsenseQA	-	0.33	88.9%
McTaco	-	-	87.1%
C3KG	-	0.2	-
GLUCOSE	-	1.6	-
System from [11]	-	0.002	77%
CommonGen	-	-	-

Table 1: Table including values for comparison measures (throughput, cost, accuracy) for all methods discussed in section 3 collected from the referenced literature. The cells that are marked with a dash, "-", suggest that the literature did not include values for the selected measures. The throughput is measured in knowledge tuples generated per day and the cost in dollars per question.

6. References

[1] Ernest Davis and Gary Marcus. Commonsense reasoning and commonsense knowledge in artificial intelligence. Commun. ACM, 58(9):92–103, aug 2015.
[2] Filip Ilievski, Alessandro Oltramari, Kaixin Ma, Bin Zhang, Deborah L. McGuinness, and Pedro Szekely. Dimensions of commonsense knowledge. Knowledge-Based Systems, 229:107347, 2021.
[3] Jonathan Gordon, Benjamin Van Durme, and Lenhart K Schubert. Evaluation of commonsense knowledge with mechanical turk.