Semantics-based Entity Summarization

Kalpa Gunaratna

Kno.e.sis, Wright State University, Dayton OH, USA
kalpagunaratna@gmail.com

Abstract. Entity descriptions encapsulate knowledge about the entities and relationships in knowledge graphs. Summarizing facts of these entities in a concise and comprehensive manner is important for consuming this knowledge. We present three entity summarization approaches covering single and multi-entity summarization.

Keywords: Entity Summarization · Ranking · Knowledge Graphs.

1 Motivation

World knowledge is captured in knowledge graphs and has gained attention in the recent past with support from the industry (e.g., Google Knowledge Graph, Bing Satori, and Amazon Product Graph) and academia (e.g., DBpedia, Yago, and Wikidata). The main building block of these graphs is the entity centric data representation either in formal languages like Resource Description Framework (RDF) or proprietary representation methods. Over time, these knowledge graphs grow in size. For example, the DBpedia English version 3.9 has 4 million entities described in 470 million RDF triples (facts) and English version 2016-04 has 6 million entities described in 1.3 billion RDF triples, averaging over 100 and 200 triples per entity, respectively.

Given the growth of entity descriptions in knowledge graphs, effective and efficient methods to consume them are required. Summarization (i.e., Entity Summarization) is a suitable method to transform this big entity-centric knowledge for efficient human or machine consumption. Entity summarization is used in Web applications and presenting a knowledge graph snippet along with text snippets is one such example (see Google or Bing search pages). Entity summarization can be categorized into two main categories based on how they are constructed: (i) extractive and (ii) non-extractive. Extractive summarization methods focus on selecting a subset of the given set of facts. Entity summarization can be further classified into two types: (i) single entity summarization and (ii) multi-entity summarization. This dissertation work supports extractive entity summaries on both single and multi-entity summarization.

2 Approaches

Single entity summarization is to summarize facts of an entity at a time. In contrast, multi-entity summarization considers multiple entities simultaneously which makes it computationally expensive and hard. This is because a selection of a fact for a specific entity summary can consequently invalidate already selected other entities' summary facts.

2.1 Single Entity Summarization

For single entity summarization, we first introduced FACES [2] (FACeted Entity Summarization) that considers both conciseness and comprehensiveness of entity summaries. In addition to creating concise descriptions, comprehensiveness influences selecting facts that cover different aspects of the entity. This is the first time that an entity summarization approach handled diversity in terms of semantics where previous efforts have been using lexical similarities in grouping facts. At the time of the introduction of FACES system, most of the entity summarization systems utilized object property facts because of the difficulties in getting semantics out of datatype properties. We extended FACES system to compute types for literals in datatype properties and utilized both object and datatype properties for summarization.

FACES We first identify groups of facts within an entity description and then use a ranking mechanism to select facts from each group that makes diversity-aware entity summaries. Note that the groups identified within an entity description are 'conceptual'. That is, facts within a group are semantically similar (at the level of abstraction) as opposed to lexically similar. For example, the grouping can discern that the two facts describing 'work place' and 'profession' agree at the abstract level of 'professional life' of a person. After grouping facts, we ranked them using a modified version of TF-IDF ranking scheme. We combined the uniqueness of a fact (property-value pair) using IDF and frequency of values using TF. The ranking function captured the essence of having popular entities and unique facts in entity summaries. The grouping approach in FACES is both incremental and semantically rich compared to all the existing methods for this purpose. Further, FACES is the first system that focused on eliminating 'semantically duplicate' facts in summaries, improving state-of-the-art results significantly.

Computing Types for Literals - FACES-E The primary FACES approach could not group datatype property facts due to the absence of semantic types in property values. The FACES-E [1] extension to FACES computed types for these literals using focus term detection and aligning these focus terms to ontology classes and entities present in knowledge graphs. These computed type information are utilized in grouping of facts in the FACES conceptual grouping algorithm and then ranked facts by extending the ranking algorithm. Literals are more unique than entities in property values and hence we extended the ranking algorithm to compute frequency of literals using the statistics of entities mentioned in the literals. There have been approaches to compute missing types and refine noisy types for entities in knowledge graphs but no technique has been proposed to improve semantic understanding of literals in knowledge graphs prior to our approach, in the literature.

2.2 Multi-Entity Summarization

Multi-entity summarization is to summarize multiple entities at the same time to collectively capture a characteristic among the entities. In this work, we emphasis relatedness of facts between the summaries. This is computationally expensive as the number of entities grow and the problem is known to be NP-hard. We proposed REMES [3] (REllatedness-based Multi Entity Summarization) that maximizes relatedness of facts between entity summaries and importance and diversity of facts within each entity summary. We adapted the quadratic knapsack problem and solved the multi-entity summarization problem using an approximation algorithm. To measure relatedness of facts, we combined entity embeddings (RDF2Vec) that measure entity-entity relatedness and term expansion based similarity on properties used in FACES grouping approach. This multi-entity summarization is a significant step, going beyond single entity summarization and have practical implications in knowledge exploration and reading comprehension applications on the Web.

3 Results

The results for FACES approach is presented in Table 1. The evaluation is based on the average overlap of the system generated summary ($Summ(e,k)$) against human generated summaries ($Summ^I(e,k)$) for the respective entities as in Equation 1. The Agreement measure is the average overlap in human generated summaries as shown in Equation 2. The gold standard created for the evaluation consists of randomly selected 50 entities from DBpedia knowledge graph [1]. The reported results for FACES are statistically significant and refer [2] for more details.

$$Quality(Summ(e,k)) = \frac{1}{n}\sum_{i=1}^{n} |Summ(e,k) \ \cap \ Summ_i^I(e,k)| \tag{1}$$

$$Agreement = \frac{2}{n(n-1)}\sum_{i=1}^{n}\sum_{j=i+1}^{n} |Summ_i^I(e,k) \cap Summ_j^I(e,k)| \tag{2}$$

The results for type computation for literals in FACES-E system is reported in Table 2. The type generation is evaluated against DBpedia Spotlight system as the baseline (by getting types of the entities identified by the baseline). Any Mean Precision (AMP) is the average 'ceiling' precision. After type computation, the summarization results against comparable systems using both object and datatype properties are presented in Table 3. We extended the gold standard used in FACES by increasing the total entities to 80. Refer [1] for more

[1] The gold standard is created due to unavailability of such datasets at the time.

System	Evaluation 1 - Gold standard					Evaluation 2 - User preference	
	k = 5			k = 10			
	Avg. Quality	FACES % ↑	Time/Entity	Avg. Quality	FACES % ↑	Study 1	Study 2
FACES	**1.4314**	NA	**0.76 sec.**	**4.3350**	NA	**84%**	**54%**
RELIN	0.4981	187 %	10.96 sec.	2.5188	72 %	NA	NA
RELINM	0.6008	138 %	11.08 sec.	3.0906	40 %	16%	16%
SUMMARUM	1.2249	17 %	NA	3.4207	27 %	NA	30%
Avg. Agreement	1.9168			4.6415			

Table 1. Evaluation of the summary quality and FACES %↑ = 100 * (FACES avg. quality - Other system's avg. quality) / (Other system's avg. quality) for k=5 and k=10, respectively, and average time taken per entity for k=5 for Evaluation 1. Evaluation 2 measures user preference % for each system. (NA stands for Not Applicable)

details. The results reported for FACES-E are statistically significant. Next, the evaluation results for REMES are presented in Table 4. The evaluation is based on a questionnaire on a Likert-scale. This is because, it is difficult to create a gold standard like FACES and FACES-E (which is already hard) for multi-entity summarization as a judge has to look at many entities at once to select facts, making the search space much bigger than the single entity scenario. The evaluation dataset consists of 10 news documents from AQUAINT and 20 documents from Wikinews datasets. Each document contains more than 2 entities tagged in them and the objective is to evaluate how systems can capture summaries for each entity considering the entire entity collection.

	Mean Precision (MP)	Any Mean Precision (AMP)	Coverage
FACES-E	0.8290	0.8829	0.8529
Baseline	0.4867	0.5825	0.5533

Table 2. Type generation evaluation. DBpedia Spotlight is used as the baseline system.

System	k = 5		k = 10	
	Avg. Quality	%↑	Avg. Quality	%↑
FACES-E	**1.5308**	–	**4.5320**	–
RELIN	0.9611	59 %	3.0988	46 %
RELINM	1.0251	49 %	3.6514	24 %
Avg. Agreement	2.1168		5.4363	

Table 3. Evaluation of the summary quality (average for 80 entities) and %↑ = 100 * (FACES-E avg. quality - Other system avg. quality) / (Other system avg. quality) for k=5 and k=10, where k is the summary length.

Q.	Wikinews					AQUAINT				
	Response: Mean (SD)			F(2,357)	LSD post-hoc	Response: Mean (SD)			F(2,147)	LSD post-hoc
	REMES	FACES	RELIN	(p-value)	(p <0.05)	REMES	FACES	RELIN	(p-value)	(p <0.05)
Q1	3.98 (1.16)	3.66 (1.08)	2.78 (1.18)	35.798 (6.772e-15)	REMES >FACES >RELIN	4.50 (0.65)	3.92 (0.92)	3.06 (1.13)	30.866 (6.427e-12)	REMES >FACES >RELIN
Q2	4.12 (0.93)	3.93 (1.08)	3.79 (1.28)	2.747 (6.500e-2)	REMES >RELIN	4.26 (1.01)	3.98 (0.89)	3.22 (1.36)	11.879 (1.700e-5)	REMES, FACES >RELIN
Q3	3.69 (0.71)	3.38 (1.41)	2.84 (0.71)	17.868 (4.022e-8)	REMES >FACES >RELIN	4.36 (0.60)	3.76 (0.96)	2.92 (1.32)	25.927 (2.267e-10)	REMES >FACES >RELIN
Q4	3.78 (1.07)	3.48 (1.07)	2.91 (1.20)	19.148 (1.260e-8)	REMES >FACES >RELIN	4.26 (0.63)	3.74 (0.80)	2.88 (1.19)	30.123 (1.086e-11)	REMES >FACES >RELIN
Q5	4.05 (0.89)	3.72 (0.91)	3.18 (1.20)	22.586 (5.805e-10)	REMES >FACES >RELIN	4.22 (0.68)	3.32 (1.04)	2.54 (1.20)	35.611 (2.447e-13)	REMES >FACES >RELIN

Table 4. Evaluating summaries using questionnaire. SD refers to standard deviation. Questions (Q) are - Q1: Summaries assisted me to get some relationships between the entities in the entity collection, Q2: The facts in each summary are diverse, Q3: The summaries helped me to better understand the document, Q4: The summaries provide me an overview of the entire entity collection, Q5: I like the summaries generated.

4 Discussion and Future Directions

The approaches presented in this dissertation research made advances in entity summarization in multiple directions. First, the FACES approach incorporated diversity in single entity summarization. Second, FACES-E approach proved it could compute types for literals and showed the usefulness of type-computed literals in creating comprehensive entity summaries. Third, REMES approach introduced a heuristic method to compute summaries for multiple entities that is tractable. The three systems presented have become state-of-the-art benchmarks in the field.

Even though the results presented outperformed existing systems, the challenges in single and multi-entity summarization are far from being solved. The systems mainly used knowledge graph based processing to rank and group facts for the summarization whereas there is wealth of knowledge in unstructured data on the Web. Getting insights from these can significantly improve the results. Further, summarization approaches can serve the purpose better if personal preferences could be identified. Hence, personalization techniques can make them more appealing to each individual user. In addition, improvements in understanding the conceptual groupings in entity descriptions can improve the quality of the summaries further.

References

1. Gunaratna, K., Thirunarayan, K., Sheth, A., Cheng, G.: Gleaning types for literals in rdf triples with application to entity summarization. In: International Semantic Web Conference. pp. 85–100. Springer (2016)
2. Gunaratna, K., Thirunarayan, K., Sheth, A.P.: Faces: Diversity-aware entity summarization using incremental hierarchical conceptual clustering. In: AAAI. pp. 116–122 (2015)
3. Gunaratna, K., Yazdavar, A.H., Thirunarayan, K., Sheth, A., Cheng, G.: Relatedness-based multi-entity summarization. In: IJCAI: proceedings of the conference. vol. 2017, p. 1060. NIH Public Access (2017)