ResearchSpace: Querying a Semantic Network

Achille Felicetti

VAST-LAB, PIN, Università degli Studi di Firenze, Italy

Querying the Graph: Fundamentals

- Divide the entities of our domain into a set of relevant **Fundamental Categories** that appear to be founded deeply in our intuitive understanding of the

- These FCs serve as domains and ranges of **Fundamental Relations**.

  - To try and cover the domain with as few FRs as possible, which a user can easily learn

  - To make some powerful distinctions that keyword search cannot do, such as discerning places from people having the same name (e.g., "*Caravaggio*")

  - To try and satisfy as many different kinds of questions as possible by asking a few more general ones, and not only the most frequently asked questions

Querying the Graph: Fundamental Classes

- **Thing = crm:E70 Thing**, comprises material and immaterial things, a special case of "what"

- **Actor = crm:E39 Actor**, comprises persons, organizations, offices, and informal groups, equal to "who"

- **Event = crm:E2 Temporal_Entity**, comprises states, historical and other periods in the sense of the CRM (crm:E4.Period), and events (crm:E5.Event) and activities (crm:E7.Activity) in the narrower sense. Can be regarded as a "when"

- **Place = crm:E53 Place**, geometric extents in space, on earth and on objects, often related to or even identified by some stable and prominent configuration of matter, such as a settlement. It is equal to "where"

- **Time = crm:E52 Time-Span**, a date-time interval, a special case of "when"

- **Concept = crm:E55.Type**, comprises all kinds of universals, such as types of things, people, events, places, species, etc. This is a special case of "what"

Querying the Graph: Theoretical Principles

- General query: a full-text search into all literals returns the associated nodes in the browser, together with minimal metadata and icons. Each node is marked by the FC it is an instance of

- For a more precise query, a user must first "select" (in the sense of the Structured Query Language (SQL) "Select" statement) the FC from which the question should return instances

- Then the user must compose a sort of "Where Clause."

  - The most simple one consists of a flat list of properties with the selected FC as domain and with range values combined by AND or OR

  - The design challenge is to find a minimal set of FRs intuitive to the user and easy to learn, which widely cover the respective discourse with high recall and a precision great enough not to be flooded by unrelated answers

Querying the Graph: Theoretical Principles

- Language disambiguates words by the relations to other words in a phrase
  - "He spoke to the museum" versus, "He walked around in the museum"
  - Seems contradictory in an ontology but not surprising for people in whatever language is translated
- "Complementary polysemy" (Pustejovsky)
  - Explained by classifying contextual expressions into relatively few, language-neutral categories
  - When a user selects a relationship term and a value, a similar mechanism to disambiguate the relationship is used to help the user:
  - The term is interpreted according to the selected FC and the FC the range value is instance of

Querying the Graph: The "from" statement

- **"From"**: a very natural relationship term describing any sort of origin or provenance

- "Things from New Guinea" (a Place)

  - things found, produced, or used in New Guinea

  - things with parts from there

  - Things produced by people coming from New Guinea.

  - Museum metadata frequently contain the term "provenance" in this sense.

- "Things from J.W. Goethe" (an Actor): different interpretation

  - things created, produced, modified, said, acquired, owned, kept, or used by him or his household

  - gifts he gave or received or awards he received.

Querying the Graph: The "from" statement

- "Things from the Parthenon" (a Thing)

  - parts or pieces of the Parthenon

  - inscriptions found on it

- "Actors (people) from New Guinea (Place)"

  - a sort of nationality concept

- "Actors (people) from Siemens Company" (Actor)

  - pertains to membership

- "Places from Time" make no sense

- All interpretations correspond to composite path expressions in the CIDOC CRM. Particular combinations of FCs as domain and range allow to find all relevant expressions in the ontology

Fundamental Relationships

- Three main categories
  1. Describing how and what something is (classification, part-whole structure)
  2. Describing what an item has undergone in its history,
  3. Describing what it may "show," say or refer to

- No relationships of intention, motivation, or cause, because they are rarely documented

Fundamental Relationships

- 14 FR identified, and among them:

  - **has type**: denotes relations of an item (i.e. any instance of a FC) to a classification, category, type, essential role, or other unary property, such as a format, material, color

  - **is part of**: denotes structural relations of an item to a wider unit it is contained in. The relationship is applicable to all FCs, except for Concept. In the case of Actors, one would rather speak of "is a member of"

  - **has met**: denotes the symmetric relation between items that were present in the same event, including time intervals and places. Applicable to any combination of FCs, except for Concepts.

Fundamental Relationships

- 14 FR identified, and among them:

  - **from**: denotes the relations of an item to a context in its history that is either significant for the item, or the item is significant for the context. "Provenance" in the widest sense, including time intervals and places

  - **refers to** or **is about**: denotes the relation of an item that is information, contains information, or has produced information to the item this information refers to or is about. The relation can even be extended to a Place from where such information originated

  - **is similar** or **the same with**: denotes the symmetric relation between items that share features or are possibly identical. It is only usual for Things to document similarity manually. There exist enough comparison algorithms that deduce degrees of similarity automatically. We do not deal with these in this work.

Domain (select)	Range(query parameter)					
	Thing	Actor	Place	Event	Time	Concept
Thing	9.has met 10.refers to or is about 11.is referred to by 13.has part 8.is similar or same with 5. from 　3.is part of 　was made from	9.has met 5.from 10.refers to or is about 11.is referred to by 12.by 　Used by 　Created by 　Modified by 　Found or 　acquired by	10.refers to 11.is referred to at 5.from 　Used at 　Created at 　Found or acquired at 　Was created 　/produced by 　person from 　Is/was located at	10.refers to 11.is referred to by 5.from 　Destroyed 　in 　Created in 　Modified in 　Used in	5.from 　Destroyed on 　Created on 　Modified on 　Used on	1.has type
Actor	9.has met 6.is creator or provider of 10. refers to 11.is referred by 4.has	3.is member of 13.has member 9. has met 5.has generator 6.is generator of 10.refers to 11.is referred by	9.has met 5.from 10.refers to 11.is referred to at	10.refers to 11.is referred to by 5.from 9.has met 　Was brought into existence at 　Was taken out of existence at 　Performed action at 　Influenced	10.refers to 5.from 9.has met 　Was brought into existence at 　Was taken out of existence at 　Performed action at 　Influenced	1.has type
Place	14.has hosted 　6.Is origin of 10.refers to or is about 11.is referred by	14.has hosted 　6.Is origin of 10.refers to or is about 11.is referred by 9.has met	3.is part or limit of 13.has part or limit	10.refers to 11.is referred by 14.has hosted	7.at 10.refers to 9.has met	1.has type
Event	6.is origin of 11.is referred by 10.refers to or is about 9.has met 　created 　destroyed 　modified 　used	12.by 11.is referred by 10.refers to or is about 9.has met 　brought into existence 　took out of existence	10.refers to or is about 11.is referred to at 7. at	10.refers to or is about 11.is referred by 13.has part 7.from	10.refers to or is about 7.at 　starts 　ends 　has duration	1.has type
Time	6.is origin of	6.is origin of	6.is origin of	6.is origin of	3.is part of 13.has part	1.has type
Concept	2.istype of	2.istype of	2.istype of	2.istype of	2.istype of	1.hastype 2.istype of

ResearchSpace

- Highly flexible and configurable Semantic Web application that uses a framework of Semantics (context and meaning) to integrate data

- Allows Semantic Search

- CIDOC CRM oriented

- Search according with different entities
  - Things, Actors, Events, Places, Events, Concepts …

- Semantic Faceting
  - Use the same contextual relationships to refine searches

●	Thing
👤	Actor
📍	Place
🕐	Time
📅	Event
💬	Concept

Thank you

- Achille Felicetti
- *VAST-LAB, PIN, Università degli Studi di Firenze, Italy*
- *achille.felicetti@pin.unifi.it*