Towards the Representation of Claims in Ontologies for the Digital Humanities

Salvatore Cristofaro,[1] **Emilio M. Sanfilippo**,[2] Pietro Sichera,[1] and Daria Spampinato[1]

[1] CNR Institute of Cognitive Sciences and Technologies (ISTC), Catania, Italy

[2] ISTC-CNR Laboratory for Applied Ontology, Trento, Italy

SWODCH, 21 September 2021

Introduction

In DH application scenarios it is often the case that:

▶ Domain knowledge/data is expressed in hypothetical terms

For example, according to Seminara, Alessandro Lanari was born in 1790 (although it is not certain...)

Approach common in the investigation of, e.g.,

▶ Phenomena for which we lack data
▶ Phenomena which require the explicit modeling of hypotheses

Introduction (cont'd)

From an ontological modeling perspective:

▶ How shall we conceive hypothetical information in an ontological coherent way?

▶ How shall we represent it?

▶ How shall we handle it through the Semantic Web?

Proposal (preliminary...), to represent explicitly

▶ Scholarly claims (including their meta-data)

▶ The dependency of (some) data on such claims

State of the art

Martin-Rodilla et al. (2019):

- ▶ Adopt fuzzy logics to express the degree of truth of data

State of the art (cont'd)

CRMinf (v.0.10.1, 2019; extension of CIDOC-CRM):

- ▶ I2 Belief
    - ▶ *J4 that* I4 Proposition Set
    - ▶ *J5 holds to be* I6 Belief Value

Example:
- ▶ | Seminara's belief | *J4t that* | Lanari was born in 1790 |
- ▶ | Seminara's belief | *J5 holds to be* | probable |

State of the art (cont'd)

CRMinf (v.0.10.1, 2019; extension of CIDOC-CRM)

- ▶ I4 Proposition Set:
    - ▶ "The effective propositions it contains should be made up of unambiguous identifiers, concepts of a formal ontology and constructs of logic" (CRMInf, p.10)

Example:

- ▶ | Seminara's belief | *J4t that* | Lanari was born in 1790 |

State of the art (cont'd)

Carriero et al.'s work about ArCo, in particular ArCo context description

Situation:

- ▶ A portion of reality satisfying a description and such that its constituting entities are related in a certain manner
- ▶ **Example:** A situation where Vincenzo Bellini is born on November 3, 1801

Interpretation:

- ▶ A situation with an epistemological nature, i.e., domain entities as they are conceived and characterized by domain experts (an "epistemological fact")
- ▶ **Example:** An interpretation where Alessandro Lanari is born in 1790 according to Seminara

Claims

Some hints:

- ▶ Result from scholarly investigations
- ▶ Accessible in an inter-subjective manner
- ▶ Possibly created collaboratively
- ▶ Depend on their creator(s)
- ▶ "Static nature" (cannot change in time)

Also,

- ▶ There can be conflicting claims about the same phenomena

Claims (cont'd)

A claim c is a description (in ArCo's sense) assigning a property p to a domain entity e such that there is a corresponding interpretation (in ArCo's sense) standing for the epistemological fact in which e satisfies p

Example:

$assignsBirthDate(c, t, p) \rightarrow$
$\exists i(hasDescription(i, c) \wedge dependsOn(i, c)\wedge$
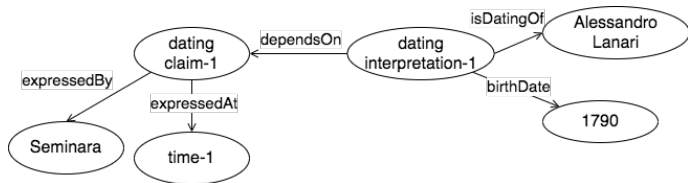$$isDatingOf(i, p) \wedge birthDate(i, t))$$

Claims (cont'd)

Figure: Example of dating claim about Lanari's birth date

Open challenges 1.

Epistemological values:

- ▶ Degree of uncertainty concerning the attribution of some properties
- ▶ **Example:** Lanari is probably born in 1790

↦ Introduction of attributes standing for "degrees of confidence" (similar to **I6 Belief Value** in CRMinf)
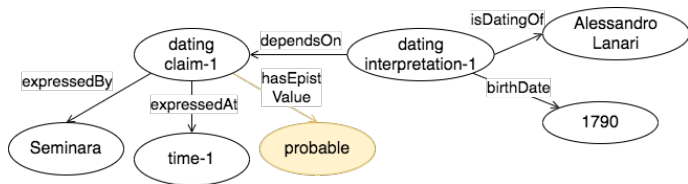
Open challenges 1. (cont'd)

Figure: Example of dating claim with epistemological value

Open challenges 2.

Conflicting claims about the same phenomena:

▶ Lanari was born in 1790 according to Seminara and in 1787 according to VIAF

$\mapsto$ "Relax" constraints?

$$bornAt(x, t, c) \land bornAt(x, t', c') \land c \neq c' \rightarrow t = t' \qquad \text{(NO!)}$$

Future steps

Further research:

- Belief modeling from both logical and ontological perspective
- Judgment aggregation theories
    - Merge compatible claims (consider different granularity levels)
    - Discard conflicting data while tracking them
- Attribution of epistemological/intentional properties to domain entities
- Robust formal approach leading to OWL modeling framework
- Test benchmark

Thank you!