

Enforcing Historical Research and Teaching with Computational Thinking and Metaphoric Abstraction

Francesco Moscato, Giuseppe Cirillo¹
DiSciPol, Second University of Naples
Viale Ellittico 31, 81100 Caserta (CE), Italy
francesco.moscato@unina2.it
¹DiSciPol, Second University of Naples
Viale Ellittico 31, 81100 Caserta (CE), Italy
giuseppe.cirillo@unina2.it

Computational Thinking (CT) is a new trend for approaching complex problems in both teaching and research activities. It requires the application of computer-science techniques in solving problems that usually belong to other knowledge and scientific domains. In particular, researchers, teachers and students may use abstraction methodologies not only to reduce complexity of problems, but also to enable reasoning by metaphors. Integration of CT with computational semantic-based reasoning, may lead to significant research results in social science and it prepares students to a new methodology for the analysis of historical and social sources. In this work, we describe a novel CT methodology based on formal semantics, metaphors and automatic reasoning, we show how it applies to the study of some concepts of modern history, and we report some results obtained by following the steps of the proposed methodology.

1. Introduction

Computational thinking [1] is related to the idea of problem-solving and to the hypothesis that thinking in a “Computational way” helps in managing very complex problems. It draws on main methodologies of the computer science and promotes their use in solving problems belonging to other domains. Someone even considers it as one of the fundamental part of the way people think and understand the world nowadays.

Computational thinking means creating and making use of different levels of abstraction, to understand and solve problems more effectively.

The idea was so promising that several enterprises (which comprises also Microsoft and Google) invested considerable sums of money in developing computational thinking methodologies for learning in primary and secondary schools [2].

Computational thinking also means thinking algorithmically and with the ability to apply mathematical concepts such as induction to develop more efficient, fair, and secure solutions. Hence, Computational Thinking research is devoted to the study of the same nature of computation [3].

The basic concepts of Computational Thinking are:

- **Decomposition** [4]: It is the basic step for the creation of algorithms following the **divide-et-impera** criterion, but in general is the ability of dividing elements of a problem in order to make easy to explain it to other persons.
- **Pattern Recognition, Generalization and Abstraction** [5,6]: these resumes the abilities of notice similarities or differences among problems components making possible prediction and shortcuts to the solution of the problem
- **Algorithm Design** [7]: this is the ability of developing a step-by-step strategy to solve the problem. Together with abstraction and generalization, it allows the definition of common, reproducible methods to solve similar problems.

In this scenario, where new reasoning methods are applied to several, different knowledge domains and areas, it seems useful the introduction of CT in teaching and analysing historical documents and concepts. Among social sciences, methodologies used for "History" learning and research still follow the "old school" approaches, that sometimes are "old" enough to be dated back to centuries ago. Even if several new analysis methods are arising in historical studies, we want to apply CT to the analysis of historical documents. The CT method we propose here resumes the steps enacted when dealing with information extraction and semantic annotation of documents. The method we present here exploits abstraction and analogies with other concepts in order to obtain information from documents and to discover semantics "hidden" to usual research methods of historical sources.

The method presented here uses Abstraction for creation of domain ontologies describing concepts in source documents and ontologies belonging to other domains that can be coped with the previous one by applying metaphoric rules, just like synonyms link several terms to the same concept.

We describe here the steps used for the application of the methodology and we apply it to a simple case study.

This work organization is the following: section 2 reports some brief motivation for the work and some related works; section 3 contains the description of the steps of the CT method, section 4 shows the application of the method to a case study. Finally, section 5 contains some concluding remarks.

2. Motivations and Related Works

Computational Thinking is relatively a new field for what application to “learning” methodologies concerns and it seem to be new at all when applied to the scientific research in social science and, in particular, in the “history” field.

Anyway, for the “Historical” point of view, an attempt to apply “transdisciplinary” case studies as a means of sustainability of learning is presented in [10]. In this work authors describe the theoretical importance of transdisciplinary case studies in learning historical science. The embryonic concepts of algorithm is hidden in that methodology that evolves step-by-step, passing through a given number of states. Transdisciplinary definition of case studies, requests:

- An ontology conceptualizing of the phenomenon/problem/case to study or to dealing with;
- An epistemology conceptualizing of the cognitive approaches the study is relying on
- A methodology for the conceptualizing of the methods the study is applying and their interaction with other methods
- A project management theory conceptualizing the approach to enact.

The same elements will be used while applying our approach, but it will use ontology conceptualization of different knowledge domain in order to merge ontologies by means metaphors. Metaphors identification will involve an conceptualization of the cognitive approaches that lead to cope together symbols from different knowledge domain. The methodology, which uses metaphors, will involve interaction with formal semantics and Semiotics domains, while the project management theory of the last point is the algorithm that is the core of the presented CT approach.

It is clear that, in addition to the traditional tools for the study of social science, Semantics (of documents) and Semiotics (a systemic relationship of signs within communication processes) have to be used.

For historical and social sciences, the approach to the text must take into account the relations between the languages and the speakers: words change their meanings through the centuries and sometimes semantics related to phrases may be lost. Researchers must take into account possible replacements of words and meanings in order to re-build original semantics from historical documents.

The concept of “meaning” is important in linguistic theories. Each meaning is defined only in relation to other meanings, within a given domain.

In addition, even if an “object” changes the original name, it continues to live linguistically. This leads to the need of paying particular attention to the use and the origins of metaphors, metonymies and other figures of speech.

Languages and semantics of documents grown more and more during the centuries in complexity. As described in [11], the semantics conceptualization of relevant concepts was first represented on a tree data

structure, but soon it became insufficient to express concepts and a complex graph (assimilated to a Labyrinth in [11]) takes its place.

In this tangle of semantic concepts and relations, we are trying to reduce complexity by using Metaphors.

In [12] the author reports an analysis of semiotic and linguistic theories coming to the fundamental assertions about the limits of semantic representations: they are too much complex to represent all the meanings of all concepts in terms of languages elements which meanings changes during centuries.

If the interpretation of the sources is demanded, in an hermeneutical perspective, to the symbols related to the words, instead of the meaning of the words, complexity decreases and possibly new "lost" meanings of documents can be retrieved.

For what information extraction and semanti-based annotation of documents, we refer to the techniques presented in [14,15], but their description is out of the scope of this work.

3. Metaphoric Abstraction Method

Metaphors are the basics elements of the methods we are going to explain.

The definition of **Metaphor**, in the Oxford Dictionary, is the following:

"A figure of speech in which a word or phrase is applied to an object or action to which it is not literally applicable",

Or:

"A thing regarded as representative or symbolic of something else"

The last definition is important because it contains the concept that a Metaphor is something having a given semantics in a domain **A**, that is used to address something else in an another domain **B**. The use, in the same document, of usual concepts and metaphors, doubtless increases the complexity of the document, but let us consider the case of a well-known, well-defined and clear domain **M**, and a set of (historical) documents, from different periods or ages, belonging to a knowledge domain **D** and that we suppose "metaphor-pruned" (i.e. without content containing metaphoric meanings).

If we are able to describe semantics of source documents in the domain **D**, in terms of the concepts of the domain **M**, regardless of the meaning (i.e. semantics) of terms in the documents, we should have the semantics of elements in the documents represented in terms of a clear set of symbols (in this case words) with defined semantics.

Reasoning on Domain **M** is simpler than reasoning on **D** and new concepts can be discovered from source documents.

More in detail, **Abstraction, Pattern Recognition and Generalization** are applied in our CT methodology.

First of all we have a set of source documents (**S**) to analyse. Let us suppose that the documents belong to different periods and that a domain ontology **D** exists able to describe the main concepts in the documents. Let us assume that terms and concepts related to them have changed during the years

so **D** is no more able to describe all the meanings of the documents since some of them have been lost due to language changes.

The first pass of the method is the retrieval of suitable Metaphors for documents contents. Let us characterize the Metaphors in ontological way and let us assume that **M** is the ontology used for defining meanings of concepts in the metaphoric domain. Both **D** and **M** have to be defined with a formal language. It is extremely important that **M** is the more possible detailed. In our example we use OWL[13] in order to define ontologies.

The next steps of the CT method are the following:

- Documents are annotated by using concepts in **D** ontology.
- The **Abstraction** technique is applied: If a metaphor is known linking concepts (and instances) from **D** to **M**, the link between concepts is formally reported. This like an operation of Ontology Merging, but **D** and **M** remain separated. At the end of this step, it is possible to refer to known concept in **D** by using a metaphoric symbol.
- The **Pattern Recognition** technique is then applied: If (graph) patterns exist in **M** but there is no related pattern in **D**, even if similar relations (in terms of graph structure) exist in the element of **D** metaphoric linked to elements in **M**, probably the metaphoric connection has been lost during the years.
- The **Generalization** technique is applied when the lost metaphoric link are re-applied to source documents, recovering lost semantics and information.

In the next section, the methodology is applied step-by-step to a simplified (but real) example.

4. Case Study

The example we propose here an example from a real case study concerning a new concept of Nobility which is different from the classical one. The results obtained in the real case study have been verified by analysing many historical documents stored in Italian National Archives (in particular inside The National Archive of Naples and The Archive of the Caserta Royal Palace) related to the period from 1800 to 1930. The identification of the document sources is out of the scope of this work. Anyway, until few years ago, It was believed that the concept of nobility was related to the concept of Ancient Family with lineage of high birth. The oldest was the origin of the familiy, the strongest was the Nobility degree it holds.

But let us apply the method previously explained to this concept.

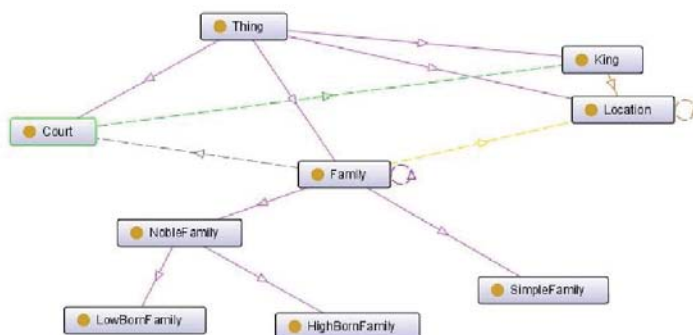


Fig.1 – Noble Family Ontology

In Fig.1 a simple ontology is reported, (the one used in the real example was more complex) describing relations between Families and The King. Families can be divided in Simple Families (i.e. no noble family) and Noble Families. These in turns can be low born or high born families. A family may descend from other families. There are Kings, that live in some locations. A location may or not be “near” another location. Families can belong to Courts and a Court follows a King.

The Ontology in Fig.1 assumes the role of the **D** ontology.

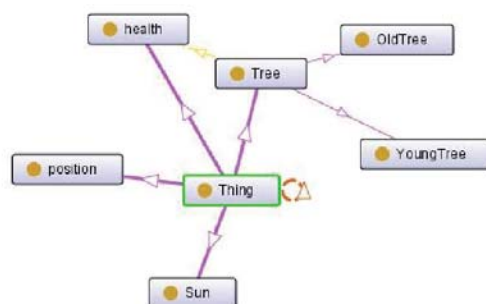


Fig.2 – Metaphor Ontology

In Fig.2, the Metaphor Ontology (it is still a simplified version of the real one) is reported. The metaphor is about trees and their proximity to sunlight. The nearest is a tree to the sunlight, the stronger it is.

There is a known metaphor between Noble Families and the Tree Ontology: the metaphor of the "family tree". It was known that the more deeper and wide was the family tree, more noble and stronger was the family.

But let us consider some Individuals in the two ontologies and let us apply reasoning on these two ontologies.

For the first ontology, let us consider the following instances:

- TheKing: a king
- CityOfTheKing: a city where the king lives
- CourtOfTheKing: the name of the court of TheKing
- FarCity: a city located far away from the CityOfTheKing
- BorgeoisFamily: a non-noble family, at the CourtOfTheKing, that lives in the CityOfTheKing
- OldHighLignageFamily1: a noble family living in FarCity and that does not belong to the CourtOfTheKing

For the second ontology, let us consider the following instances:

- MySun: A Sun
- BadHealt and GoodHealt: two kinds of tree-healts
- OldTree1: a tree with goodHealt, near MySun
- OldTree2: a tree with BadHeald, far from MySun
- YoungTree1: a tree with goodHealt, near MySun

It is clear that, in the second Ontology, A tree has good health if it is near the sun, it does not matter if it is or not old. There are obvious metaphors linking Families in the first ontology to tree in the second one. In particular, we have already reported that a strong, old, noble family is usually linked to an old tree.

From reasoning on graph structures of the two ontologies, we can assume that the Sun can be linked by metaphors to the King.

Here, after further reasoning, the lost information can be retrieved:

"The importance of family was not dictated by his lineage, but the role that covered in the courtyard, near the King".

This revelation was true especially for Vittorio Emanuele III, which used to go to Caserta for hunting, and for the family of local courtyard who acquired power to be considered powerful noble families.

5.Conclusions

In this paper we have defined a CT method for analysing by means of metaphors, historical documents. The method has been applied to a set of historical documents producing significant scientific results that have been outlined and summarized in this paper. Future works include the inclusion of automatic information extraction techniques for (semi)-automatic annotation of

source domain and the joint use of different Metaphors on different ontologies in order to improve obtained results .

Bibliography

[1] Wing, Jeannette M. "Computational thinking." *Communications of the ACM* 49.3 (2006): 33-35.

[2] Barr, Valerie, and Chris Stephenson. "Bringing computational thinking to K-12: what is Involved and what is the role of the computer science education community?." *ACM Inroads* 2.1 (2011): 48-54.

[3] Wing, Jeannette M. "Computational thinking and thinking about computing." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 366.1881 (2008): 3717-3725.

[4] Lee, Tak Yeon, et al. "CTArcade: learning computational thinking while training virtual characters through game play." *CHI'12 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2012.

[5] Ioannidou, Andri, et al. "Computational Thinking Patterns." Online Submission (2011).

[6] Basawapatna, Ashok, et al. "Recognizing computational thinking patterns." *Proceedings of the 42nd ACM technical symposium on Computer science education*. ACM, 2011.

[7] Navlakha, Saket, and Ziv Bar-Joseph. "Algorithms in nature: the convergence of systems biology and computational thinking." *Molecular systems biology* 7.1 (2011).

[8] Fensel, Dieter. *Ontologies*. Springer Berlin Heidelberg, 2001.

[9] Fellbaum, Christiane. *WordNet*. Blackwell Publishing Ltd, 1999.

[10] Scholz, Roland W., et al. "Transdisciplinary case studies as a means of sustainability learning: historical framework and theory." *International Journal of Sustainability in Higher Education* 7.3 (2006): 226-251.

[11] Eco, Umberto. *Dall'albero al labirinto: studi storici sul segno e l'interpretazione*. Bompiani, 2007.

[12] Eco, Umberto. *Trattato di semiotica generale*. Bompiani, 2002.

[13] Allemang, Dean, and James Hendler. *Semantic web for the working ontologist: effective modeling in RDFS and OWL*. Elsevier, 2011.

[14] Amato, F., Casola, V., Mazzocca, N., & Romano, S. (2013). A semantic approach for fine-grain access control of e-health documents. *Logic Journal of IGPL*, 21(4), 692-701.

[15] Amato, F., Casola, V., Mazzocca, N., & Romano, S. (2011, June). A semantic-based document processing framework: a security perspective. In *Complex, Intelligent and Software Intensive Systems (CISIS)*, 2011 International Conference on (pp. 197-202). IEEE.