Abstract

Tatsuhiko Kato[1)]
1) Japan Advanced Institute of Science and Technology,

【Key Word】Analogy, Relation, Structure-mapping theory, Vector-space models of semantics

Analogy is an ability to see the similarity among multiple objects and situations based on the “relation of relation” . Think of the situation, where you know that Taro loves Hanako, Hanako loves Mikio, and Taro is jealous of Mikio. Suppose that you came to know that Keiko loves Iwao, and Iwao loves Tomiko. Even if you do not know the situation about Keiko, Iwao and Tomiko further, you might infer that Keiko is jealous of Tomiko, because of the similarity of the situation to that of Taro, Hanako and Mikio. The ability has a critical importance for the study of cognition, because analogy is utilized in the broad spectrum of human activities, such as scientific discovery, making metaphorical expressions, education and so on. In the past theoretical and modeling studies, analogy has been studied largely in light of how people make an analogy given the object’s and situation’s “structured representations” (i.e. relations and relation of relations) in some way or the other. On the other hand, the question of “where do these relations come from?” is largely ignored. In this thesis, we try to partially answer the question, by first examining the past theorization and modeling attempts of analogy to explicate the problems that the disregard of the above question leads to. Then we formulate our own theory of analogy, and analyze the past models in terms of whether the models satisfy the requirements posed by our theory. Finally we partially test the formulation by using the past model we, although not all, satisfies the important aspects of the formulation.

First, in chapter 2, we examine the theory and models of analogy. We do this for two reasons: (1) so that we can later judge whether the models satisfy the requirements of by our own formulation, and (2) to explicate the problems which the theory and models face, when they ignore the question of where the relations come from. Specifically, we review structure-mapping theory as a representative theorization of analogy, and structure-mapping engine, Learning and Inference with Schemas and Analogy, then word2vec, as models of analogy. Structure-mapping theory (SMT) is a theory of analogy which defines analogy as making a mapping from the base domain (the knowledge you already posess) to the target domain (the objects or situations which are to be explicated), only mapping the relations (predicates which take two arguments, such as LOVE(x, y)) of the base domain, discarding the object attributes (predicates which take one argument, such as BLUE(x)). Structure-mapping engine (SME) is a model of analogy which implements the idea of SMT. The model creates the mapping from the input set of predicates and objects (base domain) to the other such input set (target domain). To create the mapping, SME follows the constraints proposed by SMT, namely only map the relations, ignoring the object attributes, and if there can be multiple choices of mapping, select the one which has the relation connected by higher-order relations (predicate of predicates). After the review, based on the relevant research, we point out that the theory and models ignore the question of where does the relation come from. For instance, SME takes the input of base and target domains as a set of predicates (relations) and objects, which are already prespecified by the modelers. The problem of this approach is that it trivializes the process of making an analogy and blurs whether the models can be said to have solved the problem, since the modelers can hand-code the input domains as much as they like. Hence, this is the problem of arbitrariness which the models face, when we ignore the question of the origin of relations.

Another set of models of semantic representation, called vector-space models of semantics (VSMs), partially overcomes the problem of arbitrariness, by making the domains of analogy from the corpus, without human hand-coding, and still being able to solve four-term analogy problems, such as “man is to woman as king is to what?” (Man :

Woman :: King : ?). We review how VSMs creates the continuous vector-space of word meanings from the corpus, especially focusing on the representative model among them, word2vec, and how word2vec solves four-term analogy problems. In addition, we mention the limitations of word2vec in terms of its ability to solve analogy problems. Overall, we conclude that although word2vec model can create the domains of analogy from the corpus without hand-coding, unlike SME, the ability of the model to make an analogy is still highly limited. In chapter 4, we amend this limitation by utilizing the method created according to our formulation of analogy.

In chapter 3, we present our formulation of analogy. We formulate analogy as the process of creating the objects and relations in base and target domains respectively, under the constraint of commutativity of functions of the relations being defined. Our formulation intends to capture the essence of analogy, namely the inference based on the "relation of relation", by defining "relation of relation" as commutativity of functions, and at the same time overcome the problem of arbitrariness, by defining analogy as creating the domains themselves. Then we introduce some concepts from category theory, such as category and functor, to concisely capture our formulation. Lastly we show our formulation can be concretely calculated in a simple analogy example.

In chapter 4, we examine SME and word2vec, in terms of whether the models satisfy the requirements posed by our formulation. We pose mainly two requirements for a model to be called to implement our formulation, namely (1) being able to construct a functor, (2) being able to construct categories, and (2a) being able to utilize a functor as a constraint to construct categories. From the examination, we observe that SME satisfies only (1), whereas word2vec satisfies (1) and (2), but not (2a). Based on this observation, we partially test our formulation on word2vec, since the model satisfies the important requirements of (1) and (2). To test our formulation, we devise a novel method to solve four-term analogy problems on word2vec based on our definition, and compare the method with other other methods, in terms of how well word2vec can solve the four-term analogy problems utilizing our method. We find the major increase in the analogy performance utilizing our method, compared to previous methods. Overall, this result of our method being able to increase the analogy performance of word2vec is a promising one for our theory.

In conclusion, in this thesis we tackled the question, in making an analogy, "where does the relation come from?". This questions is important because, without answering the question, the modelers can make arbitrary decisions when creating the inputs to the models of analogy, leading to the trivialization of analogy-making and blurring when it can be said for models to have made an analogy. We tacked this problem by defining analogy, including the process of creating the domains. Then we examined the previous models, SME and word2vec, on whether the models satisfy the requirements posed by our theory. Lastly we tested the prediction made by our theory utilizing word2vec model, and obtained the promising result for our formulation.