A *Twenty Questions* Player Using Semantic Networks

Joel Michelson, Rajagopal Venkatesaramani
Department of Computer Science
Vanderbilt University
Nashville, TN 37235

Abstract

Twenty Questions is a game that has largely been solved due to how it lends itself to crowdsourcing. In a typical game, the answerer chooses some subject, and the questioner asks yes-or-no questions which the answerer then answers. The questioners goal is to identify the subject in under twenty questions. Existing automated questioner solutions use pre-defined sets of questions to play the game and to fill in gaps in the dataset. Popular twenty questions variants available online are thus difficult to beat, as their internal databases are quite large. Using a dataset with semantic links, we attempt to create a questioner that will generate meaningful questions on its own given its knowledge of such semantic edges throughout the concept graph.

Introduction

Twenty Questions is an excellent example of a cognitive task based on semantic inference. Although the rules of the game are fairly specific, it has the potential to generalize to linguistic reasoning. Efficiently playing the game requires search space optimization, and meaningful interpretation of a semantic network. Forming questions from relational links involves rudimentary natural language processing. In order to win a game of Twenty Questions, the questioner must contain a large database of semantic information and the ability to generate questions that divide the database as optimally as it can during runtime.

The most successful existing approach uses a neural network architecture, and is online at *20q.net*. This implementation is a variant of the original 20 questions implementation that made its way into hand-held video games, produced initially by Radica in 2004, and later on licensed to Techno Source in 2011. Whereas the hand-held version had a limited subset of the website's knowledge-base, the web implementation allows dynamic expansion of the knowledge-base using information gathered both during gameplay, as well as from the internet. In contrast, our approach is an illustration of search-space optimization on semantic networks, where 20 questions is used as an evaluation metric - the problem is to minimize the number of steps taken to correctly guess what the player is thinking of using only the relations embedded in the network.

In our implementation, we use ConceptNet - a crowd-sourced weighted directed semantic network of common-sense knowledge. ConceptNet (Liu and Singh 2004) consists of a large number of common-sense relationships between words and phrases based on the Open Mind Common Sense database (Singh et al. 2002). Its data comes from many sources, including Wiktionary and Open Multilingual Wordnet (Bond and Foster 2013). Its data consists of approximately 28 million Edges between nodes and the Relations that label them. The Relations of ConceptNet lend themselves well to Twenty Questions, as they present binary properties about subjects for the questioner to use (e.g., UsedFor, IsA, PartOf, etc.). ConceptNet also supplies a weight for each edge as a rough approximation of that edges credibility. Its data is open and available in a JSON-LD API.

Model

Given the way ConceptNet is organized - multiple relations over a set of words and phrases, it is only natural that we interpret it as a graph. This graph may be interpreted either as a single layer with heterogeneity over the edges, or as a multi-layer graph, with homogenous individual layers. Though these are equivalent in terms of the information they carry, the definition of vertex degrees changes slightly in the two cases. In a single layer interpretation, it is conventional to define degrees as being edge-relation agnostic, whereas in the multi-layer model, degrees are typically defined over each layer, and convey more specific information. It is worth noting that the sum of degrees over all layers for a vertex is the relation-agnostic degree. With the intuition that degrees specific to relations may be useful, we choose to adopt the latter, a multi-layer implementation. Before building the actual graph, however, we do the following pre-processing on the dataset.

Pre-Processing

Restricting to English ConceptNet, having been constructed with Multilingual WordNet as one of its sources, contains words and relations from many major languages. In theory, each subset corresponding to a language can be separated and a multi-lingual player be formed, but we judge this to be beyond the scope of this project. We thus filter out relations that correspond to the English language. Language is denoted in each entry by the regular expression $\backslash c \backslash \langle lang \rangle$

where *lang* is a 2-letter abbreviation of the language name. The expression 'en' corresponds to English, and we use a Python script to filter these edges out.

Combining Relations Some relations in ConceptNet, though different in a lexicographic sense, may be combined for purposes of playing 20 questions. The relations *IsA* and *FormOf* for instance convey similar ideas - both are hierarchical in nature, and thus strictly one way. Similarly, we group the relations *HasFirstSubevent, HasLastSubevent* and *HasSubevent* all into *HasSubEvent*, and remove duplicate word/phrase pairs. For purposes of question generation, the order of subevent occurrence is almost always found irrelevant.

Expanding DBPedia Relations ConceptNet contains, among its diverse range of sources, information from DBPedia - a public infrastructure for semantic knowledge. DBPedia contains specific common-sense knowledge snippets as relations in a graph, which in ConceptNet have been categorized within a single relation type, with an additional string that signifies the specific relation. Examples of these are *DBPedia-Product, DBPedia-Occupation, DBPedia-Genre*. We expand each of these sub-relations into individual relations in our graph implementation.

Removing Irrelevant Relations The relation *Antonym* does not help much with question generation for a game of 20 questions. Rarely at best, if at all, is a question phrased with a negation, and given that in the English subset of ConceptNet there are 1397805 Synonym edges, as opposed to only 20041 Antonym relations, we consider it safe to discard them. Relations like *CausesDesire, DerivedFrom, EtymologicallyRelatedTo* etc. also do not provide any useful information for purposes of our game.

Constructing the Graph

We construct a multi-layered graph $G = (V, \{E_1, E_2, \ldots, E_n\})$, where edge-set E_i corresponds to the ith relation. The vertex set is common, as two words may be related to one another by more than one relation. Edges in G are directed, but we discard the weights provided by ConceptNet, as they are not truly reflective of edge-strength. In ConceptNet, edge-weight is decided by the number of times a particular word-pair was linked using a certain relation, and as it is crowd-sourced, some topics are contributed to way more than others.

Implementation

Results

Related Work

In 2005, Robin Burgener of 20qnet Inc. filed a patent for a method to play Twenty Questions using a neural network. Variants of his method are used on the 20Q website (http://20Q.net) as well as in handheld devices. In 20Q, the Questioner has access to a set number of pre-defined questions. Players may provide a number of answers besides "Yes" and "No", and each is given a different weight (e.g. "Probably" is positive with a weight of 3, and "Depends" has a weight of 0). After multiple example runs, each target-question combination is given an agreeability metric, or a weighted total of positive minus negative answers.

A neural network with two modes is trained upon this information. In the first mode, the inputs are answers to asked questions, and outputs are target objects. In the second mode, target objects are inputs and questions are outputs. During gameplay, both modes are used to find ideal questions. Initially, the first mode is used to rank target objects. The second mode is then used to find questions to provide maximal information regarding likely objects. For example, if the first network assigns high probabilities to fruits and vegetables, the question "Does it bite?" will not provide meaningful information. With this architecture, it is possible to repeatedly retrain the network using new information gathered across a number of games, at least for the online version. Such information could change the agreeability of certain target-question combinations, or it could conceivably be used to append the weight matrix with an entirely new target object.

In *The Large-Scale Structure of Semantic Networks: Statistical Analyses and a Model of Semantic Growth*, Steyvers et al. analyze the structure of three related networks: word associations, WordNet, and Roget's Thesaurus. They find that all three exhibit the qualities of small-world structures, being sparsely connected, having short average path lengths between words, and having strong local clustering. A typical node in each network is connected to very few other nodes. In WordNet, the average path length is 10.56 and the diameter is 27. Its clustering coefficient is 0.0265, which is smaller than the other examined networks, but still significantly greater than it would be for a random graph of the same size and density. Additionally, they find that the degree distributions of all three networks have power law distributions, the exponent WordNet being 3.11.

In Concept Description Vectors and the 20 Question Game, Duch et al. create a knowledge representation by creating *glosses*, or lists of keywords used to describe words. They generate CDVs from WordNet, only using words relating to animals to improve computation time.

Conclusions

Limitations and Future Work

References

Bond, F., and Foster, R. 2013. Linking and extending an open multilingual wordnet. In *ACL (1)*, 1352–1362.

Liu, H., and Singh, P. 2004. Conceptnet — a practical commonsense reasoning tool-kit. *BT Technology Journal* 22(4):211–226.

Singh, P.; Lin, T.; Mueller, E. T.; Lim, G.; Perkins, T.; and Li Zhu, W. 2002. Open mind common sense: Knowledge acquisition from the general public. In Meersman, R., and Tari, Z., eds., *On the Move to Meaningful Internet Systems 2002: CoopIS, DOA, and ODBASE*, 1223–1237. Berlin, Heidelberg: Springer Berlin Heidelberg.