Knowledge Representation and Reasoning in Answering Science Questions: A Case Study for Food Web Questions

Arindam Mitra, Chitta Baral  
Arizona State University  
{amitra7, chitta}@asu.edu

Peter Clark  
Allen Institute for Artificial Intelligence  
peterc@allenai.org

Abstract

A group of researchers from the Allen Institute of Artificial Intelligence has proposed the Aristo challenge that requires answering science questions. The goal of the challenge is to aid in the development of machines that can understand natural language, use knowledge and reason. In this work, we take a subset of those questions, namely the questions from the chapters of food web. We model a consequence operator for the food webs that given a food web and a perturbation to some of the populations aims to compute possible effects on the other populations in the food web. We then use this operator to answers questions of the kind, ‘Explain why the population of rabbits might decrease if the population of mice decreased.’ or ‘Explain why the population of rabbits might change if the population of mice decreased.’ Unlike the previous works which deal with only direct predator-prey situations, here we aim to characterize the effect(s) even when the two populations in the question are indirectly related.

Developing intelligent agents that can understand natural language, reason and use commonsense knowledge has been one of the long term goals of AI. To track the progress towards this goal, several question answering challenges have been proposed. The Winograd Schema Challenge (Levesque, Davis, and Morgenstern 2012), Allen Institute of Artificial Intelligence (AI2)’s flagship project Aristo (Clark et al. 2018), Natural Language Inference (Bowman et al. 2015), and the Stanford Question Answering Dataset, SQuAD (Rajpurkar et al. 2016) are all examples of this. Among these challenges the science question answering challenge, Aristo from AI2 is of particular interest to the community of Knowledge Representation and Reasoning (KR) for several reasons. First, answering science questions requires a wide variety of reasoning skills including qualitative, quantitative and counter factual reasoning. Second, majority of the knowledge needed to answer the questions can be found in the science textbooks. The later is an important property as it allows one to solely focus on the task of knowledge representation and reasoning without worrying about the challenging task of knowledge extraction. In this work, we focus on answering a particular type science questions, namely the food web questions. Table 1 shows two such examples.

Copyright © 2018, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Question Type 1: Explain how a perturbation leads to a certain change in a population. The first example Q1 in Table 1 is an example of a type 1 question which is taken from the 8th grade New York Science Regents exam. This type of questions looks for an explanation to an outcome given a perturbation to a food web. For example, the question Q1 in Table 1 seeks an explanation for the decreasing rabbit population. A possible answer to the question could be that “the population of rabbits decreases because the snakes would eat more rabbits”.

Note that there is also a reason for the rabbit population to increase as the competition for grass is less after the decline in mouse population. In this conflicting scenario, it is difficult to eliminate any of these two possibilities due to the absence of proper knowledge. Thus if the question asks to explain the possible effect on the rabbit population one has to list down all the possibilities. Question 2 from table 1 makes this expectation clear.

Question Type 2: Enumerate and justify all the changes in a population due to a given perturbation. Another type of question that is often asked in New York Science 8th grade exams is about explaining the possible effects on a population due to a perturbation in other populations. Question 2 in Table 1 shows an example of this kind. Unlike to the type 1 questions, this type of questions looks for all possible effects on a population. An answer to this type of questions includes exploring the possible effects and one or more justification for each scenario. For e.g. according to the solution manual provided in the Regents portal, an answer to the Question 2 in Table 1 could be as follows:

• The squid population could decrease because there would be fewer cod for them to eat.
• The squid population would increase because there would be more small animals and krill for them to eat.
• There would be less food for the elephant seals, so they would eat more squid and the squid population would decrease.
• There would be less food for the penguins, so they would eat more squid, so the squid population would decrease.
Q1. Explain why the population of rabbits might decrease if the population of mice decreased.

Q2. Explain why the number of squid might change if there were a significant decrease in the number of cod.

Table 1: Examples of food web questions.

Question Type 3 (The Consequence Operator): Can a perturbation P cause the system to go to some state S? A food web involving n populations (or species) can be in one of the 3^n states at any point of time. This is because at each time point the value of a population can either increase, decrease or remain same. However, all of those states might not be reachable due to the dependencies among the populations. Thus an important question that one can ask to test if an agent understands the concepts of the food web could be the following: Given a food web F and a perturbation P (current state), list each state s that is a consequence of P.

In this work, we aim to handle these types of questions from food web. Particularly, we have formalized the notion of a ‘good’ answer to these type of questions and a set of properties that our good answer satisfies. Finally a solver in the language of Answer Set Programming has been implemented that computes the desired answers. The solver can be downloaded from https://goo.gl/vvzcLP.

Related Work
The food web question answering task that is described in this work is an example of a qualitative question answering task. Thus the works in qualitative reasoning (Weld 1990; 1988; Forbus 1984; 2016) is also related to this work. Previous works (Nuttle et al. 2004; Takeuchi and Adachi 1983) on food web restrict themselves to only predator-prey scenario. In this work, we model the effect of a perturbation to any other population including the ones also which are not in a direct predator-prey relation.

Conclusion
The science question answering challenge has several types of questions that require sophisticated reasoning and can benefit from the KR formalisms. The domain of food web is an example of this. In this work we have developed a small theory for the food web questions. Our work here is somewhat analogous to formulations (or micro theories) of reasoning about inheritance hierarchies, and reasoning about events and actions. We believe that several such KR micro-

theories are needed to be developed and used for developing explainable question answering and text understanding systems for 8th grade science question answering. Our work here is a step in that direction.

References


