Pragmatic Axiomatization Techniques

Vinay K. Chaudhri, Richard J. Waldinger
Artificial Intelligence Center
SRI International, Menlo Park, CA 94025

Benjamin M. Good, Barbara H. Starr
Science Applications International Corporation
San Diego, CA 92121

Abstract

This paper is about domain-independent theoretical generalizations derived in the process of formalizing aspects of commonsense reasoning. One of the generalizations is a qualitative model that has a template structure, can be applied to multiple reasoning tasks. Another model reduces to a composition of linear orders, making it easy to argue its formal properties. These models allow the population of a knowledge base (KB) by non-logicians. We present experimental results on evaluation of these, and other, models by subject matter experts (SMEs).

Introduction

Commonsense reasoning can be a broad and somewhat ill defined task. Previous approaches to formalize commonsense reasoning have addressed the problem either by focusing on narrow theoretical aspects of the task [1], [2] or trying to produce all-encompassing encoding of axioms to capture common sense [3]. In this work we use a mixture of these two approaches: for a pre-specified commonsense reasoning problem specification, we encode the necessary knowledge, and in the process derive theoretical generalizations. These generalizations are aimed at obtaining better knowledge base (KB) construction techniques -- to speed up the KB construction time, and to allow non-logicians to contribute to the KB content development.

We consider the following models for commonsense reasoning: (1) a qualitative model to capture interaction between actions, (2) a model to reason with escalation of conflict in a scenario, and (3) a model to represent the capability of an agent to perform a task. These models were tested on previously unseen questions, and subject matter experts (SMEs) evaluated the answers produced. We begin by describing the problem specification. Next, we describe the models and the experimental results.

Problem Specification

The political commonsense reasoning task was defined by the *Crisis Management Challenge Problem* (CMCP) [4]. The CMCP is a collection of test questions of interest to an analyst dealing with international crises.

This research was supported by DARPA's High Performance Knowledge Bases Project.

The questions were specified using a question grammar. The grammar consisted of a set of parameterized questions or PQs, each of which had several instantiations. A sample PQ is ``What {risks, rewards} would <InternationalAgent> face/expect in <InternationalActionType>?'' An example instance of this PQ is, ``What risks Iran may expect in sponsoring terrorist attacks in Saudi Arabia?'' A sample answer would be that sponsoring terrorist attacks is a violation of international norms; therefore, sponsoring terrorist attacks may lead to sanctions against Iran.

The KB development process included several test cycles, with the test phase based on questions derived from the question grammar, but not previously seen by the system developers. A major test cycle was conducted at the end of the year with several small-scale tests in between. The results reported here are based on work conducted over a period of two years.

Teams led by Teknowledge and SAIC developed the systems for answering the questions. SRI was a part of the SAIC team. This paper concerns the work conducted primarily by SRI with assistance from SAIC.

SRI used its first-order theorem prover--**N**ew **A**utomated **R**easoning **T**oolkit, SNARK--for answering the questions [5]. The final KB contained over 20000 axioms. In this paper, we cover only salient aspects of our KB content. All axioms are accompanied by English paraphrasing of their content, and individual relation names in each axiom are not explained.

Qualitative Model of Action-Interaction

Action-interaction involved questions such as what interests may underlie an action, or what actions an agent can take to achieve an interest, or what might be the risks of performing an action.

The inference centered around two qualitative influence relations: *may-positively-influence*, and *may-negatively-influence.* We *say* that an action *A* may *positively* influence an action *B* if, by performing an action *A*, an agent would most likely further its interest in *B.* For example, by imposing sanctions against a country, an agent could further its interest in upholding international law. This relationship is purely qualitative, and does not attempt to capture the magnitude of the qualitative influence. Also, it does not imply that after performing the action *A*, the interest *B* is fully achieved. We *say* that an action *A* may *negatively* influence an interest *B*, if by performing an action *A*, an agent would most likely hurt its interest in *B.*

For example, by supporting terrorist groups in a region, an agent could hurt its interest in maintaining regional stability. Having defined the predicates for qualitative influences, the next task was to write axioms for which we developed a technique based on axiom templates.

Exploiting the Structure of a PQ

To develop the necessary KB content for a PQ, we started by generating a handful of its instances and by representing knowledge to answer them. In the process, we identified the *axiom templates* that were common to several instantiations. To illustrate this technique, let us consider the parameterized question 236 (PQ236) and a few of its instantiations.

PQ236 What {types of interest, <AnalyticalFactor>} typically underlie an <InternationalAgentType>'s decision to <InternationalActionType>?

PQ236-1. What types of interest typically underlie a country's decision to attack targets in a country?
PQ236-2. What types of interest typically underlie a country's decision to conduct a diplomatic action favoring a country?
PQ236-3. What types of interest typically underlie a terrorist group's decision to conduct a military action against a country?
PQ236-4. What economic interests typically underlie n country's decision to conduct an economic action against Iran?
PQ236-5. What types of interest typically underlie the decision of an enemy of Iran to oppose Iran's interest in construction of a major oil pipeline to Bandar Abbas, Iran from Baku, Azerbaijan through Iran?

The interests typically underlying an action are those that are positively influenced by that action[1].

```
(=>
    (and
      (may-positively-influence ?action ?interest)
      (has-mental-object-of ?interest ?agent)
      (performed-by ?action ?agent))
(typically-underlies ?interest ?action))          (A1)
```

To capture the influences between actions and interests, we designed the following template.

For all <ActionDescription> there must exist an <InterestDescription> such that <ActionDescription> positively influences it. (T1)

Here is a sample instantiation of this template that states that if a country attacks another country, it can further its interest in deterring attacks from that country[2].

```
(=>
  (and
    (performed-by ?attack1 ?country1)
    (opposing ?attack1 ?country2))
  (exists ((?deterring-aggression1 deterring-aggression))
    (and
      (interest-of ?deterring-aggression1 ?country1)
      (has-deters-against ?deterring-aggression1 ?country2)
      (may-positively-influence ?attack 1
                    ?deterring-aggression1)))))     (A1)
```

From the above instances of PQ236, PQ236-1 through PQ236-4 is amenable to solution by this axiom template. PQ236-5 is context dependent and requires the system to have knowledge about the economic benefits of pipelines and the desire to oppose enemies. To cover the greatest possible question space, we focused on those instantiations that were amenable to solution by an axiom template.

Definition of an Axiom Template

An axiom template lets us generate rules with a special structure. Axiom template is a rule with place holders for plugging in conjunctions of literals. Sometimes, an implication in the opposite direction may exist with a minor rearrangement of quantifiers over the same set of literals. It should be possible to systematically generate the sets of literals to be plugged into the template. For example, they may correspond to a set of objects that exist in the domain of application.

For example, for the axiom template T1, <ActionDescription> and <InterestDescription> are the place holders. In its instantiation A2, <ActionDescription> is replaced by the conjunction of (performed-by ?attack ?country1) and (opposing ?attack ?country2).

Example Applications of Axiom Template

Actions leading to interests:
PQ237. What types of international action does a desire toward <InterestType> typically lead to?

The notion of ``typically leads to'' can be formalized as

```
(=>
  (may-positively-influence ?action ?interest)
  (typically-leads-to ?interest ?action))     (A3)
```

The influence between actions and interests needed for this question is slightly different from T1.

If an agent has an <InterestDescription> then there exists an <ActionDescription> such that the <ActionDescription> positively influences the <InterestDescription>. (T2)

[1] We use the ANSI KIF syntax for writing the axioms [6]. We also follow the convention of naming variable after types, for example, ?action is a variable of type action.

[2] The notation (?deterring-aggression1 deterring-aggression) in the quantifier means that ?deterring-aggression is a variable of type deterring-aggression.

```
(=>
  (and
    (interest-of ?deterring-aggression1 ?country1)
    (has-deters-against ?deterring-aggression1 ?country2))
 (exists ((?attack1 attack))
    (and
      (may-positively-influence ?attack 1
            ?deterring-aggression1))
      (performed-by ?attack1 ?country)
      (opposing ?attack1 ?country2))))                (A4)
```

The ?attack1 in the axiom A4 is a hypothetical attack, and the axiom does not imply that the country will necessarily perform the attack. To specify actions that actually occur, we use the occurs-in relation, for example (occurs-in ?attack1 persian-gulf-war) would imply that the ?attack1 actually occurred during the Persian Gulf War.

Axiom A4 can be obtained from axiom A2 by swapping the description of the action in the antecedent of A2 with the description of the interest in the consequent of A2 and adjusting the quantifiers accordingly. Variations of PQ237 required lengthier chains of inference. For example:

What types of international action does a desire toward decrease in the price of oil typically lead to?

The basis for the solution of this question is the inverse relationship between production and price. We represented this as follows. According to the laws of supply and demand (other things being unchanged): if the supply of a product increases, the price of that product will decrease.

```
(=>
 (and
     (has-product-measure ?product
      ?economic-measure-of-product)
     (has-production-capacity ?agent
      ?economic-measure-of-product)
     (has-price ?product ?money)
     (has-action-effect ?action
       ?economic-measure-of-product ?increase-effect))
   (has-action-effect ?action ?money ?decreasing-effect)))
```

A similar axiom may be written to indicate that if the supply decreases, the price will increase. Next, we link this to our qualitative model of influences using an axiom that states that if an action achieves a certain effect, that action positively influences the interest in achieving that effect.

```
(=>
 (and
     (has-action-effect ?action ?thing ?effect)
     (has-mental-object-of ?interest ?agent)
     (performed-by ?action ?agent)
     (has-interest-in-thing ?interest ?effect))
 (may-positively-influence ?action ?interest))
```

Using these axioms, we could infer that to increase the price of oil an agent would typically try to decrease its production and vice versa.

Action responses:

Most of our reasoning with interests and actions has been on the premise that the agents perform actions to further one or more of their interests. The responses often depend on the situation as in the following PQ.

PQ219. <ContextSpec>, what actions might <InternationalAgent1> take [with respect to <InternationalAgent2>] [in response to <EventSpec>]?

PQ219 makes the reasoning more specific by setting a context and by giving a specific event that needs to be responded to. We can apply our qualitative model of influences by first identifying the interests of an agent in a situation, then checking which of those are negatively influenced by the given action, and then identifying how those negative influences could be positively influenced. This intuition is captured in the following axiom.

If ?action1 performed-by ?agent2 negatively influences ?interest1 of ?agent1, and ?action2 positively influences ?interest1, then ?action2 is a viable response to ?action1 with respect to ?agent2.

```
(=>
    (and
     (may-negatively-influence ?action1 ?interest1)
     (has-mental-object-of ?interest1 ?agent1)
     (may-positively-influence ?action2 ?interest1)
     (not (equal ?action1 ?action2))
     (performed-by ?action1 ?agent2))
    (and
     (may-respond-to-action ?agent1 ?action1 ?action2)
     (starts-after-starting-of ?action2 ?action1)
     (has-with-respect-to-agent ?action2 ?agent2)))
```

Benefits and risks:

PQ39. [{During, After} <TimeInterval>], what {risks, rewards} would <InternationalAgent> face/expect in <InternationalActionType>?

To assess the benefits and risks of performing an action, we developed a cause-effect model. This model is based on five predicates: (1) *cause-event-event* to represent the effects that are *definitely* caused by an action, (2) *may-cause* to represent the effects that *may* be caused by an action, (3) *may-prevent* to represent actions that *may be prevented* by an action, (4) *maleficiary* to represent the negative effects, and (5) *beneficiary* to represent the positive effects.

The following axiom was based on the proceedings of a Paris ministerial conference on terrorism cited in the source material for the CMCP specification [4]. One of the resolutions at the conference was to punish terrorist acts.

```
(=>
 (performed-by ?terrorist-attack ?agent)
 (exists
   ((?punishment punishment ))
```

```
(and
  (may-cause ?terrorist-attack ?punishment)
  (maleficiary ?punishment ?agent)
  (object-acted-on ?punishment ?agent)))
```

This axiom is an instance of the following axiom template

Forall ?action with description <ActionDescription1> there exists ?action2 with <ActionDescription2> such that ?action1 causes ?action2, and there is a maleficiary or beneficiary relation between ?action2 and the performer of ?action1. (T3)

If an action may cause another action, and that action benefits the doer of the first action, then that second action is a benefit of the first action

```
(=>
  (and
    (may-cause ?action1 ?action2)
    (performed-by ?action1 ?agent)
    (beneficiary ?action2 ?agent))
 (benefit-of-action  ?action1 ?action2 ?agent))
```

A similar axiom may be written for risks by using the maleficiary relation.

Modeling Escalation[3]

The escalation reasoning task involves determining, for an event in a given scenario, whether the event leads to escalation or de-escalation of conflict, or is retaliation for some earlier event. The task was specified using the following PQ.

PQ210. <ContextSpec>, is <EventSpec1> an {{escalation, de-escalation} of conflict, retaliation for <EventSpec2>}?

Intuitively, an action in a context is an escalation if it is in response to another, less hostile action. Conversely, an action is de-escalation if it is in response to a more hostile action. Finally, an action is retaliation for another if it is a reaction to the other and the two actions are opposed to each other or have contrary interests.

The formal representation of the scenario consists of a series of event descriptions. Each event description included the causal relationship between events, using which we formalized retaliation as follows.

```
(=>
(and
  (performed-by ?action1 ?agent1)
  (performed-by ?action2 ?agent2)
  (opposing ?action1 ?agent2)
  (opposing ?action2 ?agent1)
  (occurs-in ?action2 ?context)
  (cause-event-event ?action1 ?action2))
(retaliation ?action2 ?action1 ?context)))
```

To infer whether an event leads to escalation or de-escalation, we associated a hostility level with each action. Using the hostility level, we defined escalation as follows. An ?action2 is an escalation in ?context if it is in response to an ?action1 and is of greater hostility level than ?action1.

```
(=>
  (and
    (occurs-in ?action2 ?context)
    (cause-event-event ?action1 ?action2)
    (greater-hostility ?action2 ?action1))
   (escalation ?action2 ?context))
```

A similar axiom for de-escalation can be written. Let us now explain the process of associating hostility levels with events. Herman Kahn ("On Escalation") proposed a notion of hostility level based on a 44-stage linear scale [9]. This idea was modified by a CMCP SME during an interview. Our own scale currently has three components: damage level, weapon level, and proximity level.

The damage level is an estimate of what kind of damage the action involves. A military attack, for example, is likely to involve population damage, which is the most severe level. Public criticism of one government by another is likely to reach only the verbal damage level, which is much less severe.

The weapon level reflects the kind of weapons the attack involves. For example, an attack using biological weapons is most severe. An attack that involves no weapons has the least severity on the weapon scale.

The proximity level concerns the location of the attack. An attack on the heart of another country has the most severe proximity level; an attack that is outside the borders of the target country has the lowest proximity level.

Hostility levels of two actions can be compared with a "lexicographic ordering." More precisely, the action with the higher damage level has the higher hostility level; if the damage levels are equal, the action with the higher weapon level has the higher hostility level; if the damage level and the weapon level are equal, the action with the higher proximity level has the higher hostility level.

Figure 1, created by Kestrel Institute, illustrates how the computation of escalation can be compositionally constructed from the theory of linear orders. Constructing each of the damage, weapon, and proximity scales amounts to renaming the sort names in the theory of linear orders. Such renaming is consistency preserving. The lexicographic ordering of the three scales is a co-product that also is consistency preserving. Using the hostility model in conjunction with the escalation axiom is a definitional extension, which is consistency preserving as well. Whenever we can reduce a reasoning task to a well-

[3] This was joint work between SRI International and Kestrel Institute.

understood formal theory, it greatly eases the task of ensuring its formal properties.

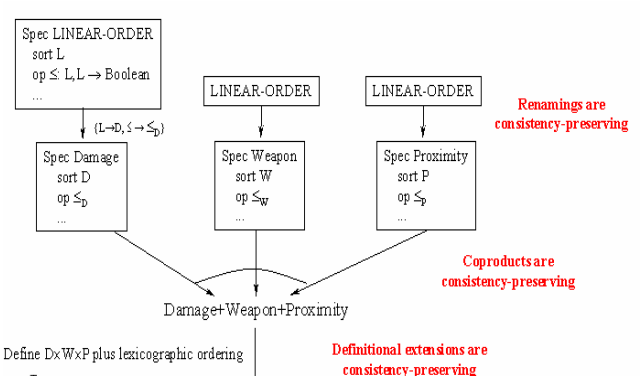

Figure 1. Compositional construction of the hostility model from the theory of linear orders.

Modeling Capability

Several CMCP questions required reasoning about capabilities of an agent to perform a certain task. For example,

PQ35. Can <InternationalAgent1> sponsor <InternationalActionType> inside <InternationalAgent2>'s borders?

PQ76. Does <InternationalAgent> have a <MilitaryOrganizationType> that can perform <InternationalActionType>?

Many of the instantiations of PQ35 involved the ability of an agent to sponsor terrorist attacks. Instantiations of PQ76 involved actions that were military in nature. The capability of an agent to perform military actions is mostly determined by whether it possesses the necessary resources. The capability for performing terrorist attacks is dependent not just on resources, but on the ideology of a state and the past history of doing such acts. For example, the following axiom states that a state can perform terrorist attacks if it has supported such attacks in the past.

```
(=>
    (and
     (performed-by ?supporting-action ?country)
     (object-acted-on ?supporting-action ?terrorist-group))
    (capable-of-because ?country supporting-terrorist-attack
                        ?supporting-action)))
```

A country may perform an air strike if it possesses the requisite aircraft and if the target is within its range.

```
(=>
    (and
     (possesses ?country ?aircraft-class)
     (=<
      (in-km (distance-between ?country ?place))
      (in-km (range ?aircraft-class)))
     (object-found-in-location ?thing ?place))
   (exists ((?air-attack air-attack))
    (and
       (capable-of ?country ?air-attack)
       (performed-by ?air-attack ?country)
       (device-used ?air-attack ?aircraft-class)
       (event-occurs-at ?air-attack ?place)))
```

It is quite common to state that a country is capable of performing an action, even though it is usually a specific group or organization that actually performs the action. The following axiom captures this intuition. If an agent possesses an organization with a certain capability, the agent also has that capability.

```
(=>
   (and
     (possesses ?agent ?organization)
     (capable-of ?organization ?action))
   (capable-of ?agent ?action))
```

Experimental Results

We tested our KB that contained the axioms defined in this paper. The KB, of course, contained many other axioms. For populating the KB using templates, a table was constructed showing actions on one side and interests on the other side. The SMEs were asked to indicate positive and negative influence between the actions and interests. For the escalation model, SMEs were asked to linearly order various weapon, damage, and proximity levels.

The test questions were not previously seen by the developers of the KB, but derived from the question grammar that was known in advance. SMEs, who had degrees in political science, graded the answers. The criteria for scoring included correctness, recording the sources, and explanation. For the purpose of this paper, the correctness of the answer produced is the most relevant measure, and is the only score reported.

The first column of Table 1 specifies the reasoning task tested. The second column specifies the number of test questions. The sample sizes are small because the SME time needed to generate and grade the answers was quite expensive. The third column specifies the questions with non-zero scores. The scores were given on a scale from 0 to 3. The last column shows the average correctness score for the questions with non-zero scores.

We can see that the axioms for capability, and interests leading to actions worked well at least for the questions on which they were tested. For questions involving interests underlying an action, the average score was 2 (or 66.67%); for questions involving benefits and risks, it was 2.2 (or 73%). The main reason for the less-than-perfect score was lack of precision in the answer: either too many answers were returned or the answers did not take into account the details of the situation. For example, one of the questions asked about the interests underlying sanctions imposed

against an ally. As an answer to this question, the SMEs expected the system to take into account the fact that the sanctions are being imposed against an ally; the axiom template did not capture that fact.

Table 1. Correctness scores for reasoning tasks

Reasoning Task	Number of Test Questions	Questions with non-zero correctness score	Average correctness score
Interests underlying an action	3	3	66.67%
Interests leading to actions	2	2	100.00%
Action responses	4	2	50%
Benefit/Risk	8	6	73%
Escalation	4	3	77%
Capability	5	4	100%

The average score for the question involving action responses was 1.5 (or 50%). The axiom driving the reasoning "If ?action1 performed-by ?agent2 negatively influences ?interest1 of ?agent1, and ?action2 positively influences ?interest1, then ?action2 is a viable response to ?action1 with respect to ?agent2" did not meet the SME approval. For example, the imposition of U.S. sanctions on China would damage Chinese economic interests, whereas a trade agreement would support China's economic interests. But a Chinese trade initiative to the U.S. would not be "a viable response" to the U.S. sanctions.

The average score for reasoning with escalation was 2.3 (or 77%). For the case that had less than the perfect score, the error was in the KB encoding that gave the incorrect inference. The inferences produced by the model were accurate.

Future Work

The main improvement in the axiom template technique is to enrich the template by adding more detail, for example, by specifying the conditions under which each of the influences holds. Qualitative reasoning models provide techniques for representing the conditions under which the influences hold [7],[8], which can bring the necessary precision in the axiom templates. The escalation model involving needs to be generalized when the escalation is computed using a 64-stage model [9]. Past history and possessing necessary devices are only two of several factors that determine the capability of an agent to perform a task. The capability model needs to be enriched to take into account other kinds of requisites for performing a task.

Summary

We considered three models for commonsense reasoning: (1) a qualitative model to capture interaction between actions and interests, and between actions and actions, (2) a model to reason with escalation of conflict in a scenario, and (3) a model to represent the capability of an agent to perform a task. The qualitative model involving actions followed a template structure making it easy for the non-logicians to populate the KB. The escalation model reduced to a composition of linear orders. The capability model made use of possession and past history to reason about an agent's ability to perform a task.

Our claim is not that the axioms reported here provide a complete theoretical characterization of commonsense notions such as typicality, escalation, etc. Instead, the claim is pragmatic that when SMEs tested the axioms reported here on previously unseen questions, they exhibited fairly high accuracy. Furthermore, these axioms could be written fast by non-logicians either by instantiating axiom templates or by specifying linear orders that were designed by people trained in logic. The results are grounded in a practical commonsense reasoning task: political commonsense. The generalizations of the results are, however, applicable across the domains.

References:

[1] J. McCarthy. 1968. "Programs with Common Sense," in *Semantic Information Processing*, M. Minsky, Ed. Cambridge, Mass.: The MIT Press, 403-418.

[2] J. R. Hobbs and R. C. Moore. 1985. *Formal Theories of Commonsense World*. Norwood, New Jer., Ablex Publishing Corporation.

[3] D. B. Lenat and R. V. Guha. 1990. "Building Large Knowledge-Based Systems: Representation and Inference in the CYC Project," 336.

[4] IET, Alphatech, P. Cohen, and P. S. Research. 1997. "HPKB Year 1 End-to-End Challenge Problem Specification, Version 1.1," December . See http://www.iet.com/Projects/HPKB.

[5] M. Stickel, R. Waldinger, M. Lowry, T. Pressburger, and I. Underwood. 1994. "Deductive Composition of Astronomical Software from Subroutine Libraries," in *Proceedings of the Twelfth International Conference on Automated Deduction (CADE-12)*, June, 341--355.

[6] M. R. Genesereth and R. E. Fikes. 1999. "Knowledge Interchange Format, Reference Manual."

[7] K. Forbus. 1984. "Qualitative Process Theory." Artificial Intelligence, 24:85--168.

[8] J. Rickel and B. Porter. 1999. "Automated Model Composition." *Artificial Intelligence Journal.*

[9] J. Kingston, 1999. Knowledge Acquisition for Crisis Management: Interests and Actions, Technical Report, University of Edinburgh. See http://www.ai.sri.com/hpkb/intact.html.