Representing diagnosis knowledge

David Poole

Department of Computer Science, University of British Columbia, Vancouver, B.C., Canada V6T 1Z4, and Canadian Institute for Advanced Research

This paper considers the representation problem: namely how to go from an abstract problem to a formal representation of the problem. We consider this for two conceptions of logic-based diagnosis, namely abductive and consistency-based diagnosis. We show how to represent diagnostic problems that can be conceptualised causally in each of the frameworks, and show that both representations of the same problems give the same answers. This is a local transformation that allows for an expressive (albeit propositional) language for giving the constraints on what symptoms and causes can coexist, including non-strict causation. This non-strict causation can be represented in each framework without adding special reasoning constructs to either framework. This is presented as a starting point for a study of the representation problem in diagnosis, rather than as an end in itself.

1. Introduction

This paper defines an abstract "knowledge representation" problem and considers the problem of representing knowledge in the context of diagnostic systems. We consider two diagnostic formalisms and compare how we can represent a domain in each so that each representation produces the same answer.

This paper contains many of the results of [14], recast in light of latter developments. We take a different perspective from subsequent (to [14]) papers [2, 3, 9], in that we consider the problem of going from an abstract problem to a representation of the problem rather than the problem of just going from one representation to another. While the local transformation methods may work for simple theories, there is still much to be learnt about what needs to go into any axiomatisation [16], and the mappings are not so straightforward.

One of the ideas that we are tackling is to represent subtle distinctions in the domain with rather weak representation languages. One of the main reasons for pushing weak representation languages is that we can see what they can and cannot represent, and only complicate the representations when necessary. In this paper we consider how to represent causal relations that are not strict implications (e.g., a cold may cause sneezing, but it does not imply sneezing). There are no new non-strict implications in either representation language we consider, but they can both represent strict and non-strict causes.

Like Console et al., [2, 3] and unlike Konolige [9] we consider acyclic causal structures (some c cannot cause itself). Acyclicity allows us to have a local transfor-
mation from the domain knowledge to the representations unlike the global transformations of Konolige (see [9, section 5.2]).

This paper does not contain the final answer to this problem; there is still much that has to be understood about representing more complex problems than that considered here [16].

1.1. THE KNOWLEDGE REPRESENTATION PROBLEM

DEFINITION 1.1

Given a formalism (formal language plus an inference relation), the knowledge representation problem is the problem of going from a problem \( P \) to a representation \( R_P \) of \( P \) in the formal language so that the use of the inference relation for the representation will yield a solution to the problem. \[ \]

In this definition, a problem is "a question raised for inquiry, consideration, or solution" (definition from Webster's Ninth New Collegiate Dictionary). This is not a formal representation of the problem. Many problems can be conceptualized in different ways, and these different conceptualizations may have different representations (even for the same formalism) and may have very different computational and ergonomic properties.

Definition 1.1 is depicted in fig. 1. We want to define the knowledge representation (KR), the inference relation and the interpretation for the answers so that this diagram commutes\(^1\).

This notion of knowledge representation should be contrasted with the view of knowledge representation (KR) research as defining and analysing formalisms, without the knowledge representation (as defined here) being explicit (see e.g., much of the work on nonmonotonic reasoning [8]). The knowledge representation

\[^1\] A diagram commutes if each directed path to the same point produces that same answer. In this case, the solution to the problem obtained by going via the representation and computation is the same as the solution obtained going directly from the problem to the solution.