Maintaining Meaning of Information when it is Shared amongst Information Systems

Janet Aisbett, The University of Newcastle, Australia
Shuxin Zhao, The University of Newcastle, Australia
Greg Gibbon, The University of Newcastle, Australia
C. N. G. Dampney, The University of Newcastle, Australia

ABSTRACT

Continuity of healthcare refers to the ability for multiple points of care to deal seamlessly with a patient. Tools for improving continuity of care, such as patient Electronic Health Records, need to draw on information from diverse information systems. However, reliable sharing of information across systems is known to be a difficult problem. An approach to information integration that is possible when there is a controlling organisation or a consensus amongst partners is to enforce compliance of new information systems with a master model. This requires a feasible and reliable way to check whether the data structures of a target information system conform to those of the master system. A method based on the mathematical category theory has been previously proposed. This paper describes the underlying model and investigates a stricter definition of compliance that could be employed with Electronic Health Records. It then outlines implementation in a semi-automated compliance checking system.

Keywords: healthcare systems; knowledge and information management

INTRODUCTION

An Electronic Health Record is envisaged as a collection of event summaries which contain those details pertinent to a person’s future healthcare. Electronic health records require information to be drawn from diverse data sets managed by different health-related organisations. The recent decisions to implement electronic health record systems have, therefore,
added impetus to work directed at improving the reliability of exchange of information between healthcare information systems.

Despite a huge volume of research on schema integration over the last decade or so (e.g., Rahm & Bernstein, 2001), automated translation between different systems is still a practical challenge. A technically simpler approach is to enforce compliance of information systems with a master model, that is, with a high level abstraction of the relevant domain. Here, to be compliant means that the data structures of the candidate information system are in some sense consistent with those of the master system.

A formal notion of compliance and a method of compliance checking was developed by Dampney and collaborators (e.g., Colomb, Dampney, & Johnson, 2001; Dampney, Johnson, & Munro, 1992; Dampney, Pegler, & Johnson, 2001) based on constructs from mathematical category theory. Modelling using category theory makes explicit structure that may be carried implicitly in the constructs of more complex description languages and, thus, enables information structures to be directly compared.

This paper reviews some pertinent schema matching methods, then what the category theory approach of Dampney et al. brought to such representational work. It then discusses a stronger form of compliance, originally presented in Aisbett, Dampney, and Gibbon (2004). Implementation of this compliance checking approach involves preliminary pruning and matching steps, which the paper outlines.

BACKGROUND

Concepts are associated with descriptors or attributes, and sets of attribute values are often treated as synonymous with a concept instance. A variety of diagrammatic schemes have been developed for modelling concepts and the relationships between them. In these schemas, concepts are classified into one of a few categories, elements of which are depicted in a fixed way (e.g., as a box or an oval) which may however vary between schemes. In object-oriented schemes, concepts may be passive or active. Relationships between concepts are depicted as lines connecting the concept elements or entities. Since relationships are also classified into several categories, the lines are adorned to indicate their type, for example, as an “is-a” or a “part-of” relationship. Cardinality of a relationship may also be shown though adornment of the connecting line, that is, as a one-to-many or many-to-many relationship. Auxiliary to the diagram, or possibly shown on it, are various constraints which reflect the fact that only some states of the system may be valid. These include constraints on the allowable values of attributes.

Because of the practical dominance of the relational model, much early schema integration work was in terms of entities, attributes, and constraint relationships. For example, schemas have been modelled as triples of sets, namely a set of classes (concepts), a set of attributes, and a set of integrity constraints limiting the allowable states (e.g., Vidal & Winslett, 1994). Attributes can be interpreted as functions between classes and may be either value attributes mapping to classes such as height, or referential attributes mapping into other domain classes. In addition to the