Formal Properties of Conditional Independence in Different Calculi of AI[1]

Milan Studený

Institute of Information Theory and Automation
Academy of Sciences of Czech Republic
Pod vodárenskou věží 4, 18208 Prague 8, Czech Republic
e-mail: studeny@cspgas11.bitnet

Abstract. In this paper formal properties of CI in different frameworks are studied. The first part is devoted to the comparison of three different frameworks for study CI: probability theory, theory of relational databases and Spohn's theory of ordinal conditional functions. Although CI–models arising in these frameworks are very similar (they satisfy semigraphoid axioms) we give examples showing that their formal properties still differ (each other). On the other hand, we find that (within each of these frameworks) there exists no finite complete axiomatic characterization of CI–models by finding an infinite set of sound inference rules (the same in all three frameworks). In the second part further frameworks for CI are discussed: Dempster–Shafer theory, possibility theory and (general) Shenoy's theory of valuation–based systems.

1 Introduction

The concept of *conditional independence* (CI) seems to attract attention of researches in last two decades. The properties of CI were studied in several branches of AI, let us mention some of them :

- probabilistic reasoning
- theory of relational databases
- theory of ordinal conditional functions
- theory of belief functions
- possibility theory.

The reason is evident – the knowledge concerning independence (resp. dependence) can simplify many reasoning tasks. Here, we give a short survey of current state.

So far, the most advanced results concerning properties of CI were achieved in the probabilistic framework. The concept of stochastic CI has been studied in probability theory and modern statistics for many years (see [2, 11]). Its importance for the theory of probabilistic ES consists in the fact that any CI–statement can be interpreted

[1]This research was supported by the internal grant n.275105 of the Academy of Sciences of Czech Republic "Conditional independence properties in uncertainty processing".

as certain qualitative relationship among symptoms. This brings the possibility of finding the adequate structural model of such ES. This role of CI was discerned and highlighted by the group around J. Pearl [6] (A.Paz, D.Geiger, T.Verma) but many other researchers dealt more or less explicitly with this concept.

Nevertheless, an analogy of the concept of CI was studied (even earlier) is the theory of relational databases [7]. The counterpart of CI in that theory is the concept of *embedded multivalued dependency* (EMVD). Note that an equivalent concept of *qualitative conditional independence* also appeared in ES theory [8].

Another framework in which the concept of CI appeared, is Spohn's theory of *ordinal conditional functions* [12]. This theory, motivated from philosophical point of view, gives a tool for a mathematical description of the dynamic handling of deterministic epistemology, in this sense it constitutes a counterpart of the probabilistic description of epistemic state. As soon as the concept of CI for ordinal conditional functions was introduced, researchers began to study its properties [4], especially for the special class of natural conditional functions (NCF) called "disbelief function" in [9] or "ranking function" in [3].

One of the most popular approaches to dealing with uncertainty in ES is Dempster–Shafer's theory of belief functions. The concept of CI for variables on which belief functions are defined (i.e. the parallel with concepts of CI studied in probability theory and in the theory of NCFs) was introduced lately by Shenoy [10]. Another definition of (unconditional) independence (also for variables) appeared in [1].

Further framework in which CI can be studied is Zadeh's *possibility theory.* This theory was formulated in the end of seventies as certain model of qualitative description of subjective judgements. Lately, Shenoy introduced the concept of CI also in this field. The above mentioned Shenoy's work [10] gives certain unifying point of view on different calculi dealing with CI. He introduced very abstract concept of *valuation–based system* (VBS) and defined CI for VBSs.

2 Basic Definitions

All above mentioned frameworks for study CI have some common setting. Throughout the paper a collection of nonempty finite sets $\{X_i\,;\ i \in N\}$ is supposed to be given. The index set N is also nonempty and finite. Whenever $\emptyset \neq A \subset N$ the symbol X_A denotes the cartesian product $\prod_{i\in A} X_i$. Power set of a set S will be denoted by $exp\, S$.

The symbol $T(N)$ is reserved for the collection of triplets $\langle A, B, C\rangle$ of pairwise disjoint subsets of N where $A \neq \emptyset \neq B$. Following Pearl [6] we call every subset of $T(N)$ *dependency model over* N. A dependency model is called *semigraphoid* iff it is closed under following four inference rules (called axioms by many authors):

$\langle A, B, C\rangle \rightarrow \langle B, A, C\rangle$ *symmetry*
$\langle A, B\cup C, D\rangle \rightarrow \langle A, C, D\rangle$ *decomposition*
$\langle A, B\cup C, D\rangle \rightarrow \langle A, B, C\cup D\rangle$ *weak union*
$[\,\langle A, C, D\rangle \,\&\, \langle A, B, C\cup D\rangle\,] \rightarrow \langle A, B\cup C, D\rangle$ *contraction.*

Now, we show how dependency models arise in probability theory, the theory of relational databases and the theory of conditional functions.