Journal of Mathematical Psychology 82 (2018) 73–83

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

journal homepage: www.elsevier.com/locate/jmp

Are quantum-like Bayesian networks more powerful than classical Bayesian networks?☆

Catarina Moreira *, Andreas Wichert

Instituto Superior Técnico, INESC-ID, Av. Professor Cavaco Silva, 2744-016 Porto Salvo, Portugal

ARTICLE INFO

Article history:
Received 26 June 2017
Received in revised form 4 November 2017

Keywords:
Latent Variables
Quantum cognition
Bayesian networks
Quantum-like models

ABSTRACT

Recent works in the literature have proposed quantum-like Bayesian networks as an alternative decision model to make predictions in scenarios with high levels of uncertainty. Despite its promising capabilities, there is still some resistance in the literature concerning the advantages of these quantum-like models towards classical ones.

In this work, we developed a Classical Latent Bayesian network model and we compared it against its quantum counterpart: the quantum-like Bayesian network. The comparison was done using a well known Prisoner's Dilemma game experiment from Shafir and Tversky (1992), in which the classical axioms of probability theory are violated during a decision, and consequently the game cannot be simulated by pure classical models. In the end, we concluded that it is possible to simulate these violations using the Classical Latent Variable model, but with an exponential increase in its complexity. Moreover, this classical model cannot predict both *observed* and *unobserved* conditions from Shafir and Tversky (1992) experiments. The quantum-like model, on the other hand, is shown to be able to accommodate both situations for *observed* and *unobserved* events in a single model, making it more suitable and more general for these types of decision problems.

1. Introduction

The task of determining human judgments under uncertainty has got increasing attention in the scientific literature in the last decade (Moreira & Wichert, 2016b). More specifically, several models that are capable to predict or explain human decisions that are inconsistent with the laws of classical probability theory and logic (Crosson, 1999; Kuhberger, Komunska, & Josef, 2001; Lambdin & Burdsal, 2007; Tversky & Shafir, 1992) have been recently proposed. These models turn to quantum probability to explain human decision-making and are part of a new emerging discipline called Quantum Cognition (Busemeyer, 2015; Wang, Busemeyer, Atmanspacher, & Pothos, 2013).

Recent research shows that quantum-based probabilistic models are able to explain and predict scenarios that cannot be explained by pure classical models (Bruza, Wang, & Busemeyer, 2015; Busemeyer & Wang, 2015). However, there is still a big resistance in the scientific literature to accept these quantum-based models. Many researchers believe that one can model scenarios that violate the laws of probability and logic through classical probabilistic decision models that are often used in machine learning (Murphy, 2012). These violations of the laws of probability theory are hard to explain through classical theory and can have different types: violations to the Sure Thing Principle (Savage, 1954), disjunction/conjunction errors (Tversky & Kahneman, 1983), Ellsberg (Ellsberg, 1961)/Allais (Allais, 1953) paradoxes, order effects (Sudman & Bradburn, 1974), etc.

To accommodate these violations, several quantum-like models have been proposed in the literature. Note that, the term *quantum-like* is simply the designation that it is used to refer to any model that is applied in the domains outside of physics and that use the mathematical formalisms of quantum mechanics, abstracting them from any physical meaning and interpretations.

Although, the quantum cognition field is recent in the literature, there have been several different quantum-like models proposed in the literature. These models range from dynamical models (Busemeyer, Wang, & Lambert-Mogiliansky, 2009; Busemeyer, Wang, & Townsend, 2006; Pothos & Busemeyer, 2009), which make use of unitary operators to describe the time evolution since a participant is given a problem (or asked a question), until he/she makes a decision, to models that are based on contextual probabilities (Aerts & Aerts, 1994; Khrennikov, 2009b; Yukalov & Sornette, 2011). Quantum-like dynamical models have

☆ This work was supported by national funds through Fundação para a Ciência e a Tecnologia (FCT) with reference UID/CEC/50021/2013 and through the PhD grant SFRH/BD/92391/2013. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

* Corresponding author.
E-mail addresses: catarina.p.moreira@tecnico.ulisboa.pt (C. Moreira), andreas.wichert@tecnico.ulisboa.pt (A. Wichert).

https://doi.org/10.1016/j.jmp.2017.11.003
0022-2496/

also been proposed in the literature to accommodate violations to the Prisoner's Dilemma Game (Pothos & Busemeyer, 2009), study the evolution of the interaction of economical agents in markets (Haven & Khrennikov, 2013; Khrennikov, 2009a) or even to specify a formal description of dynamics of epigenetic states of cells interacting with an environment (Asano et al., 2013). On the other hand, quantum-like models based on contextual probabilities, explore the application of complex probability amplitudes to define contexts that can interfere with the decision-maker (Khrennikov, 2005b, 2009b, 0000). For a survey about the applications of quantum-like models for the Sure Thing Principle, the reader can refer to Moreira and Wichert (2016b).

In the literature, it is clear and acceptable that simple and pure probabilistic models cannot accommodate human decisions that violate the laws of classical probability theory and logic (Busemeyer & Bruza, 2012). But can a more complex classical model simulate the paradoxical findings reported in the literature? In order to answer this question, we propose the application of latent variables in classical models to accommodate these paradoxical findings. By latent variables, we mean random variables that are *hidden*, that is, they cannot be directly measured in an experimental setting, but can be indirectly inferred from experimental data. These variables bring great advantages to cognitive models, because many observed variables can be condensed into a smaller number of hidden variables, enabling a dimensionality reduction of the model. For instance, in Psychology or Social Sciences, one can use latent variables to summarise the influence of several variables, such as beliefs, personality, social attitudes, etc., over the concept of *h* uman behaviour (Bollen, 2002; Griffiths, Steyvers, & Tenenbaum, 2007).

A well-known classical model that can include such dependencies is the Bayesian network (Pearl, 1988). This model represents relationships between random variables (such as causal and conditional dependencies) in an acyclic probabilistic graphical structure. Bayesian networks are powerful inference models that have been successfully applied over the years in different fields of the literature, mainly in artificial intelligence, genetics, medical decision-making, economics, etc.

In this work, we developed a classical Bayesian network that makes use of *latent variables* and we compared it against its quantum counterpart, the quantum-like Bayesian network, which was previously proposed in Moreira and Wichert (2016a). In the end, we conclude that it is possible to simulate the violations to the Sure Thing Principle using the classical Bayesian network with latent variables with an exponential increase in its complexity, however this model cannot predict both *observed* and *unobserved* experimental conditions from Shafir and Tversky (1992). On the other hand, the quantum-like model is shown to be able to accommodate both situations for *observed* and *unobserved* events in a single and general model. Note that the Sure Thing Principle is a concept widely used in game theory and was originally introduced by Savage (1954). This principle is fundamental in Bayesian probability theory and states that if one prefers action A over B under state of the world X, and if one also prefers A over B under the complementary state of the world X, then one should always prefer action A over B even when the state of the world is unspecified.

This manuscript is organised as follows. In Section 2, we introduce a general definition for latent variables. In Section 3, we present the prisoner's dilemma game and several works of the literature that report experiments, which violate the Sure Thing Principle in this game. In Section 4, we propose a classical Bayesian network model that makes use of Latent Variables to accommodate the paradoxical findings of the prisoner's dilemma game. In Section 5, it is introduced the quantum-like Bayesian network proposed in the work of Moreira and Wichert (2016a) as an alternative model to accommodate the several paradoxical findings reported in the literature. In Section 6, we make a discussion about the complexity involved in exact probabilistic inferences over classical and quantum-like Bayesian networks. We end this work with Section 7, which summarises the main points of this work: that the quantum-like Bayesian network model poses advantages towards the classical model with latent variables, since it can simulate both *observed* and *unobserved* phenomena in a single network, in contrast with the classical model requires extra hidden nodes (contributing to a decrease in efficiency) and cannot accommodate both *observed* and *unobserved* experimental conditions in a single model.

2. Latent variables

Most of the times, the data that is recorded (or observed) does not provide all the information that is needed to model a decision scenario. In these situations, latent variables are used to model complex patterns that we do not have the complete data for.

There is not a general and formal definition for latent variables (Bollen, 2002). Since it is a concept that is widely used across different multidisciplinary areas, it can be defined differently according to its application. However, a very simple and informal definition can be given as variables that are not directly observed from data, but can be inferred using the information of the variables that were recorded (Anandkumar, Hsu, Javanmard, & Kakade, 0000). Instead of specifying concrete relationships between variables, latent variables enable the abstraction of these relationships allowing a more general representation, which can be inferred from the observed variables.

In this work, we will use latent variables in a probabilistic graphical model, more specifically in a Bayesian network. Generally speaking, a Bayesian network is an acyclic probabilistic graphical model, which provides an intuitive way of specifying probabilistic relationships and dependencies between random variables (Griffiths et al., 2007). These relationships are specified through a joint distribution over the set of all random variables in the model, and each node specifies conditional dependencies over its parent nodes. Under this representation, a random variable becomes latent when it is unobserved (or unknown), which suggests a *local independence* definition, according to Bollen (2002). When a latent variable is constant (for instance, a prior probability representing a person's cognitive bias towards some topic), the observed variables become independent. More formally, the independence between random variables and the latent variables is given by Eq. (1).

$$Pr(X_1, X_2, \ldots, X_n) = Pr(X_1|h)Pr(X_2|h)\cdots Pr(X_n|h). \tag{1}$$

Given a set of observed random variables $X_1, X_2, \ldots, X_n$ and some vector of latent (hidden) variables h, the joint probability $Pr(X_1, X_2, \ldots, X_n)$ corresponds to the product of the conditional probabilities of each random variable X_i over the associated latent variable, $Pr(X_1|h)Pr(X_2|h)\cdots Pr(X_n|h)$.

Consider Fig. 1. Suppose you have a parameterised acyclic probabilistic graphical model over the parameter ϕ. We will assume that node H represents a latent variable, because it is not directly observed (or it is hidden) for some given reason: it might be too expensive to collect its data, it might have been not recorded or we simply might not have access to the process generating the observed data.

Given a dataset of collected data D of size M, the above network consists in a tuple $\langle h[m], x[m]\rangle$, where h is parameterised instance of the latent variable H and x an instance of the random variable X. The likelihood (a measure similar to a probability, which provides