Submission ID# 11906

Bayesian Causal Maps as Decision Aids in Venture Capital Decision Making: Methods and Applications [†]

Short title:
Bayesian Causal Maps for VCs

BENEDICT KEMMERER
University of Kansas
1300 Sunnyside Avenue
Lawrence, KS 66045-7585
Tel: (785) 864-7318
Fax: (785) 864-5328
E-mail: bkemmerer@ku.edu

SANJAY MISHRA
University of Kansas
1300 Sunnyside Avenue
Lawrence, KS 66045-7585
Tel: (785) 864-7528
Fax: (785) 864-5328
E-mail: smishra@ku.edu

PRAKASH P. SHENOY
University of Kansas
1300 Sunnyside Avenue
Lawrence, KS 66045-7585
Tel: (785) 864-7551
Fax: (785) 864-5328
E-mail: pshenoy@ku.edu

Submitted to the Entrepreneurship Division (ENT) Division

[†] We would like to express our gratitude to the expert venture capitalist for his committed participation in this study. We also would like to thank the attendants of the Management Science & Technology Seminar at the University of Kansas for their insightful comments.

ABSTRACT

Improving venture capitalists' decision processes is key to reducing failure rates for venture capital backed companies and to improving portfolio returns. In this paper we describe the use of a novel technique—Bayesian causal maps—to support and improve venture capital decision making. We combine causal mapping and Bayesian network techniques to construct a Bayesian causal map. The resulting probabilistic model represents salient features of decision makers' mental models and inference processes. Heeding the call of prior research, we focus not on generating incremental descriptive insights into the actual decision processes of venture capitalists, but concentrate on creating a model designed to serve as a practical decision aid. The process of constructing Bayesian causal maps is presented using the real case of an experienced venture capitalist specializing in early stage, high technology investments. We then turn our attention from the construction of Bayesian causal maps to their application. Bayesian causal maps can support venture capital decision making through bias reduction, reduction of unsystematic error, assumption surfacing, what-if analyses, and by facilitating systematic learning from experience, both individual and collaborative. We discuss the advantages of Bayesian causal maps as well as the limitations and challenges inherent in their construction and use. Suggestions for future research are also offered.

Keywords:

Cognitive Theory (1) - Cognitive mapping (10) - Decision making (12) - Venture capital (49)

Characterization:

Empirical

Does not specifically deal with women's entrepreneurship

By most measures, including available capital, number of funded proposals, deal size, number of venture capital firms, and number of venture capitalists (VCs), the venture capital industry has grown rapidly in the last decade. According to figures by the National Venture Capital Association (NVCA), US deal volume has expanded from investments in 1,317 companies totaling $ 3.4 billion in 1990 to investments in 5,458 companies worth $ 103.8 billion in 2000. At the same time, a large number of venture capital backed companies continue to fail, with at least 40% of backed ventures failing to provide a return (Ruhnka, Feldman, & Dean, 1992). Consequently, reducing this failure rate by even small amounts should have a substantial impact on venture portfolio returns (Zacharakis & Meyer, 1998).

One key to reducing VC-backed venture failure rates is to understand and improve venture capitalists' decision processes. Thus, a large number of studies have attempted to investigate how venture capitalists make decisions (e.g., Tyebjee & Bruno, 1984; MacMillan et al., 1985, 1987; Sandberg et al., 1988; Hisrich & Jakowicz, 1990; Hall & Hofer, 1993; Zacharakis & Meyer, 1995; Muzyka et al., 1996; Zacharakis & Meyer, 1998; Shepherd, 1999; Zacharakis & Meyer, 2000; Zacharakis & Shepherd, 2001; Shepherd & Zacharakis, 2002).

Counter to much of this literature, we do not focus primarily on the description of VC decision making, but rather on the creation of a decision aid that can support and improve decision making over time. We concur with Shepherd and Zacharakis (2002) who argue that decision aids represent a crucial next step in the research agenda on VC decision making. Consequently, while the accurate modeling of key features of a venture capitalist's mental models and decision process is a necessary secondary objective of this paper, the primary objective is to build a model that supports decision making processes in addition to describing them.

Our paper is structured as follows. First, we briefly provide some background on decision models and decision aids. Second, we employ a novel technique—Bayesian causal maps—to develop a probabilistic decision model for an experienced venture capitalist. Bayesian causal maps (Nadkarni & Shenoy, 2001), a new causal mapping based form of causal Bayesian networks, combine techniques from the causal mapping and Bayesian Network literatures. We provide some theoretical background for both causal maps and Bayesian networks and outline a methodology for constructing Bayesian causal maps. In the first real-world application of Bayesian causal maps, we apply this methodology to the decision model of an experienced venture capitalist specializing in early-stage, high-tech investments. Third, we describe applications for the resulting model in decision making and decision process improvement. Fourth, we discuss the advantages as well as the challenges and limitations inherent in using Bayesian causal maps for decision modeling. Finally, suggestions for future research are offered.

DECISION MODELS AND DECISION AIDS

Decision Models

Decision models based on policy capturing or judgement analysis have a long history in the social science literature (e.g., Hoffman, 1960; Naylor & Wherry, 1965). One of the rationales for research into mental models in general, and decision models in particular, is the fact that simplification is prevalent in decision making (e.g., Newell & Simon, 1972; Schwenk, 1984a). Boundedly rational decision makers (Simon, 1955; Cyert & March, 1963) make decisions based on simplified models of their environment and most people use only three to seven informational cues at a time (Stewart, 1988). Thus, while quantitative models will never fully capture all the variables and intuitive decision rules of an expert decision makers, there is some suggestion that they can adequately represent the most salient features of the simplified decision processes boundedly rational humans engage in.

Decision models can be distinguished along several dimensions. One is the method of elicitation. Studies investigating venture capital decisions have used, for example, participant direct report (e.g., Tyebjee & Bruno, 1984; MacMillan et al., 1985, 1987), verbal protocols (Sandberg et al., 1988; Hall & Hofer, 1993; Zacharakis & Meyer, 1995), repertory grids (Hisrich & Jakowicz, 1990), and conjoint analysis (e.g., Muzyka et al., 1996; Zacharakis & Meyer, 1998; Shepherd, 1999; Shepherd & Zacharakis, 2002).

Based on Brunswik's lens model anchored in social judgement theory (Brunswick, 1955, 1956), Shepherd and Zacharakis (2002) distinguish between *actuarial models,* which attempt to optimally weight those cues that predict new venture *performance*, and *bootstrapping models,* which aim to capture the VC's assessment policy, i.e. the cues that predict VC investment *decisions.*

Other model categorization schemes include compensatory (where an unacceptable value on one criterion can be offset by a high value along another one) vs. non-compensatory models (e.g., Riquelme & Rickards, 1992), and black-box input-output models vs. models that focus on the actual decision *process* (e.g., Sandberg et. al, 1988). In addition, models can be characterized by the extent to which decision variables as well as independence relationships and interaction terms are specified a-priori, as is the case for example in many conjoint analyses (e.g., Shepherd & Zacharakis, 2002).

Our Bayesian causal map model is a bootstrapping model, reflecting the *decision* model of one particular venture capitalist. Consequently, it predicts venture *performance* only to the extent that the participant VC is indeed an expert. Also, this means that the model is not representative of other VCs, venture capital firms, types and stages of investments. However, bootstrapping models outperform the vast majority of VCs on whose decision process they are based and are pragmatically much more feasible than actuarial models (Shepherd & Zacharakis, 2002). Pragmatic considerations are important and legitimate considerations in the development of decision aids. The benefits of our Bayesian causal map approach to model building will be discussed later in greater detail, but the model can incorporate both compensatory and non-compensatory relationships and imposes very few a priori restrictions in independence or the nature of interactions.

We would like to reiterate at this point that our focus is not on building a prescriptive or even universally applicable descriptive model, but to create a decision aid that can help the VC on whose expertise it is based and also supports collaborative learning and expertise sharing.

Decision Aids

Decision aids are methods that attempt to improve decision making. They include a wide variety of tools and techniques, including devils advocacy (e.g., Schwenk, 1984b) and simulation (Hall & Menzies, 1983) for example. Our discussion will focus on formal, quantitative decision models which also can function as decision aids (Zacharakis & Meyer, 1998; Zacharakis & Meyer, 2000; Shepherd & Zacharakis, 2002).

Bootstrapping models potentially improve decision accuracy by reducing unsystematic error and by reducing the likelihood that individual informational cues are over- or underweighted due to their idiosyncratic salience in a particular decision situation. Research from a variety of areas has shown that bootstrapping models often outperform decision makers (cf. Zacharakis & Meyer, 2000; Shepherd & Zacharakis, 2002 for further details). In addition, the model building process in itself can also reduce decision making bias. Hodgkinson and colleagues (1999) report that the mere process of creating a causal map reduced decision bias (framing bias) in a sample of both undergraduate students and senior managers.

More importantly, decision models can enhance systematic learning. The high failure rates of venture capital backed businesses demonstrate both the need and the potential for improvement. Decision models in general, and Bayesian causal maps in particular, support learning in multiple ways.

First, assumptions and decision rules are made explicit in the model building process. This is a form of cognitive feedback (Shepherd & Zacharakis, 2002), providing VCs with insights into their cognitive processes. Thus, decision rules move from being tacit, taken-for-granted assumptions to visually and quantitatively explicated models that can be actively and consciously scrutinized by decision makers or expert coaches. There is evidence that VCs have imperfect insight into their own decision process (Zacharakis & Meyer, 1998). The explication and scrutiny of beliefs is a key first step in belief updating and revision. Bayesian causal maps also allow for easy what-if and scenario analyses providing VCs with a deeper understanding of the nature and implications of their decision rules.

Second, models can be used to systematically evaluate outcome feedback (e.g., Shepherd & Zacharakis, 2002) . Once the outcome of a venture can be observed, the initial assumptions about the states of the cue variables can be compared to actual outcomes and "wrong" decisions can be traced to specific failures in either (1) the cue variable state assessment (e.g., management quality was overestimated), (2) the overall decision rules (e.g., the model did not consider a key variable or interdependency, or did not weigh its impact correctly) , or (3) the implementation (marketing strategy was correctly assessed as good, but implementation of the marketing campaign was mishandled). As failures in variable assessment or inference processes become clear, decision rules can be revised. However, it has been pointed out (Gupta & Sapienza, 1992; Shepherd & Zacharakis, 2002) that there is no way of knowing what venture performance would have been without the influence of the VC him or herself. In addition, performance depends on a multitude of exogenous variables that cannot possibly be incorporated into the model, making attribution and learning more difficult.

Third, formalized decision models enable collaborative learning. Models can be compared between novices and experts, or between partners of the same firm. Due to their explicit nature they can serve as a point of reference and departure for both individualized coaching as well as group learning in a peer or within-organization context.

In the following section, we will describe a specific decision aid—Bayesian causal maps—including its theoretical foundations, associated methodology, and practical application.

BUILDING A BAYESIAN CAUSAL MAP

Bayesian causal maps incorporate techniques both from the causal mapping and from Bayesian network arena. Before we outline and implement the methodology, we provide a brief theoretical background on each.

Causal Maps

The use of causal maps to represent individuals' mental models is commonly traced to Tolman (1948) and was popularized in the social sciences by Axelrod (1976). He defined a causal map as "a specific way of representing a person's assertions about some limited domain such as a policy problem. It is designed to capture the structure of the person's causal assertions and to generate the consequences that follow from the assertions." (Axelrod, 1976, p. 72). Causal maps capture knowledge, expertise, and assumptions in the form of directed cause-effect and means-end relationships between variable-like concepts. A sample map is depicted in Figure 1.

Insert Figure 1 about here

Assumptions. The use of causal maps in decision analysis rests on three core assumptions about the role of cognition in decision-making. (1) Causal associations are a key way in which decision problems can be described and understood (Huff, 1990); (2) revealed causal maps represent to a significant extent the actual mental models of the decision maker; and (3) these simplified, causality-based representations of the decision environment form the basis for decision making and managerial action. While we cannot present a full discussion here (e.g., Axelrod, 1976; Huff, 1990) there is significant evidence that these assumptions are not unreasonable. There is, for example, evidence for the congruence of external with private statements (Fiol, 1995) as well as the congruence of causal maps with later action (Bonham and Shapiro, 1976).

Applications. Causal maps have been used to study a wide variety of phenomena, including strategic change (Barr, 1998; Barr and Huff, 1997), environmental adaptation (Barr et al., 1992; Fahey and Narayanan, 1989), joint venture formation (Fiol, 1989; 1990), software operations support expertise (Nelson et al., 200), and intrapreneurship (Russell, 1999). Another stream of research pioneered by Eden (e.g., Eden, 1991; Eden and Ackermann 1993; 1998) has used causal maps to define the actual decision problem and to reveal hidden decision premises.

Derivation. Causal maps can be derived based on documentary sources (e.g., letters to shareholders, speech and interview transcripts) as well as using more intrusive measures such as card sorts or repertory grids (e.g., Bougon, Weick and Binkhorst, 1979). Causal maps can also be developed directly by, or in consultation with, the research participant, aided by software such as *Decision Explorer* (Banxia, 2000).

Analysis. Analysis of causal maps has largely focused on the content and the structure of causal maps. The *content* of maps is usually analyzed largely qualitatively or using a combination of qualitative and quantitative techniques (e.g., in the strategic change and environmental adaptation studies cited above). In order to investigate the structural properties of causal maps, researchers have investigated and operationalized constructs such as comprehensiveness (e.g., Langfield-Smith and Lewis, 1989), density (e.g., Calori et al., 1994; Eden et al., 1981; Laukkanen, 1994), and centrality (e.g., Eden, Jones and Sims, 1983; Fiol and Huff, 1992). The dynamic properties of causal maps, i.e. the direct analysis of implied decision outcomes have received scant attention in the management literature (see Axelrod, 1976 for exceptions from the political science arena).

Bayesian Networks

Bayesian networks have their roots in attempts to represent expert knowledge in domains where expert knowledge is uncertain, ambiguous, and/or incomplete. Bayesian networks are based on probability theory. A primer on Bayesian networks can be found in the work of Spiegelhalter and colleagues (Spiegelhalter et al., 1993).

A Bayesian network model is represented at two levels, the qualitative and the quantitative level. At the qualitative level, we have a directed acyclic graph in which nodes represent variables, and directed arcs describe the conditional independence relations embedded in the model. Figure 2 shows a Bayes net consisting of four discrete variables: Management Market Know-how (M), Management Quality (Q), Potential Revenue (R), and Decision to Invest (D). At the quantitative level, the dependence relations are expressed in terms of conditional probability distributions for each variable in the network. Each variable X has a set of possible values called its *state space* that consists of mutually exclusive and exhaustive values of the variable. In Figure 2, for example, Management Market Know-how, Management Quality, and Potential Revenue have two states each: 'High' and 'Low.' Decision to Invest has two states: 'Go' and 'No go.' If there is an arc pointing from X to Y, we say X is a *parent* of Y. For each variable, we need to specify a table of conditional probability distributions, one for each configuration of states of its parents. Figure 2 shows these tables of conditional distributions—P(M), P(Q | M), P(R), and P(D | Q, R).

Insert Figure 2 about here

Semantics of Bayes nets. A fundamental assumption of a Bayesian network (or Bayes net for short) is that when we multiply the conditionals for each variable, we get the joint probability distribution for all variables in the network. This assumption is equivalent to assuming conditional independence relations in the joint probability distribution. These conditional independence assumptions can be read directly from the Bayesian network graph as follows. Missing arcs (from a node to its successors in the sequence) signify conditional independence assumptions. Thus, the lack of an arc from M to R signifies that M is independent of R; the lack of an arc from Q to R signifies that Q is independent of R; and the lack of an arc from M to D signifies that D is conditionally independent of M given Q and R. Pearl (1988) and Lauritzen et al. (1990) describe other equivalent graphical methods for identifying conditional independence assumptions embedded in a Bayesian network graph.

Conditional independence and causality. Unlike a causal map, the arcs in a Bayesian network do not necessarily imply causality. The (lack of) arcs represent conditional independence assumptions. How are conditional independence and causality related? Conditional independence can be understood in terms of relevance. If Z is conditionally independent of X given Y, then this statement can be interpreted as follows. If the true state of Y is known, then in assigning probabilities to states of Z, the states of X are irrelevant. In practice, the notion of direct causality is often used to make judgments of conditional independence. Consider a situation where X directly causes Y and Y in turn directly causes Z, i.e., the causal effect of X on Z is completely mediated by Y. Then it is clear that although X is relevant to Z, if we know the true state of Y, further knowledge of X is irrelevant (for assigning probabilities) to Z, i.e., Z is conditionally independent of X given Y. This situation is represented by the Bayesian network

$X \rightarrow Y \rightarrow Z$ in which there is no arc from X to Z. As another example, consider the situation where X directly causes Y and X also directly causes Z. Although knowledge of Y is relevant to Z (if Y is true then it is more likely that X is true which in turn means that it is more likely that Z is true), once we know the true state of X, then further knowledge of Y is irrelevant to Z, i.e., Y is conditionally independent of Z given X. This situation is represented by the Bayesian network $Z \leftarrow X \rightarrow Y$ in which there is no arc from Y to Z or vice-versa. Finally as a third example, consider the situation where X and Y are two independent direct causes of Z, i.e., X and Y are unconditionally independent. But if we learn something about the true state of Z, then X and Y are no longer irrelevant to each other (if Z is believed to be true and X is false, then it is more likely that Y is true), i.e., Y is not conditionally independent of X given Z. This situation is represented by the Bayesian net $X \rightarrow Z \leftarrow Y$ in which there is no arc from X to Y or vice-versa.

Making probabilistic inferences. Inference (also called probabilistic inference) in a Bayesian network is based on the notion of *evidence propagation*. Evidence propagation refers to an efficient computation of marginal probabilities of variables of interest, conditional on arbitrary configurations of other variables, which constitute the observed evidence (Spiegelhalter et al., 1993). Once a Bayesian network is constructed, it can be used to make inferences about the variables in the model. The conditionals given in a Bayesian network representation specify the *prior* joint distribution of the variables, e.g., the base line investment probability, given no information about the investment opportunity. If we observe (or learn about) the values of some variables (e.g., quality of management), then a *posterior* joint distribution of the variables can be calculated, including updated probabilities for key variables of interest (e.g., invest decision). Thus, the joint distribution of variables changes each time we learn new information about the variables.

A Bayesian Causal Map for Venture Capital Decision Making

Below, we elaborate the process of constructing a causal map-based Bayesian network of an expert venture capitalist, implementing for the first time a methodology proposed by Nadkarni and Shenoy (2001). The procedure consists of five steps: (1) Elicitation of a raw causal map from an expert venture capitalist, (2) preprocessing of the raw causal map, (3) assignment of states to variables, (4) assignment of probabilities to states, and (5) refinement and validation.

Elicitation of raw causal map. The first step in building the raw causal map was to interview somebody with expert knowledge of the domain, in this case an experienced venture capitalist. By using a semi-structured interview, the concepts in the map are allowed to emerge from the data, rather than being predetermined a priori by the researcher (Carley and Palmquist, 1992). In order to elicit the Venture Capitalist's mental model, we not only asked him to give a general description of the venture capital investment process, but also asked him to describe three real world cases to us. These cases involved both successful and unsuccessful investments, as well as a case where the venture capital firm decided not to make an investment after studying the venture. Using case studies is much less intrusive than directly asking experts to explicate a general model of their decision processes. Experts often have a limited understanding of their own decision models (e.g., Zacharakis and Meyer, 1998). Thus, a study of *revealed* decision criteria through case descriptions usefully complements the direct elicitation of decision models.

The interview lasted 1.5 hours. The interview was transcribed, yielding in excess of 13000 words. A revealed causal map was created from this textual base, by (1) identifying causal statements in the text, and (2) constructing a raw causal map with cause concepts linked to effect concepts by arrows indicating the direction of causality. The raw causal map is shown in Figure 1.

Causal statements are assertions that can be represented in the format "the more A, the more/less B" (Axelrod, 1976, p. 258). A concept can therefore be thought of as a variable that is able to take on different values. The following is an example of a causal statement:

> "[…] with the name that he had in the pharmaceutical industry. So that was a real asset, that was one of the reasons we made the investment".

A causal statement thus identified is then transformed into a graphical representation in the form of a raw causal map. The map consists of vertices, linked by arrows. The vertices represent the concepts. The relationship between concepts is represented by an arrow in the direction from the cause to the effect concept. The causal relationship is either positive or negative, i.e., the cause has either a promoting or retarding influence on the effect variable (Axelrod, 1976, p. 10f.). This is represented in the form of a "+" or "–" above the arrow. An example appears in Figure 3. The totality of all (domain relevant) causal statements in a given text forms a raw causal map[1].

---

[1] For further illustrations and descriptions of how to derive revealed causal maps (including from non-textual sources) see the works by Axelrod (1976), Huff (1990), Eden and Spender (1998), and Nelson et al. (2000).

Insert Figure 3 about here

The example above was an extreme case of an *explicit* causal statement. However, semi-structured interviews and other texts also contain a multitude of *implicit* causal statements. It was here, where we slightly deviated from standard causal mapping methodology. While the coding of implicit causal statements is standard practice (cf. Wrightson, 1976), researchers generally favor a conservative approach to including questionable causal statements. In our case, there were two reasons that let us err on the side of inclusion: (1) As we will discuss below, in Bayesian networks, counter to causal maps, the absence of linkages has meaning (conditional independence), thus it is important to represent all linkages present. (2) The methodology for constructing Bayesian causal maps necessarily includes a feedback process, where the original interviewee is asked to verify the map and certain linkages, something that is often not possible in the case of many causal mapping studies (based on textual sources) who thus need to be more conservative.

Preprocessing of the causal map. Before probability values can be assigned to the causal map, the raw causal map has to be preprocessed in order for it to be compatible with Bayesian network theory (Nadkarni and Shenoy, 2001). The four necessary steps are discussed below.

Conditional independence. Causal networks can either be modeled as dependence maps (D-maps) or independence maps (I-maps) (Pearl, 1988). In a D-map, concepts (variables, vertices) connected by arrows are indeed dependent. However, concepts lacking a direct connection may or may not be conditionally independent. In contrast, in an I-map, the absence of a connection between two concepts does indeed imply conditional independence, given the state of other variables in the map, but the presence of a link may or may not imply dependence. Causal maps, by their nature are D-maps, while Bayes nets are I-maps. Thus, the raw causal map needs to be transformed into an I-map, usually in consultation with the expert on whom the raw causal map is based. Due to the number of nodes in the map, grid-based techniques were infeasible, so we asked the research participant to audit the map directly for missing links, focusing initially on thematic sub-maps to keep the task manageable. At the end of the procedure, the Bayes net was both a D-map and an I-map, i.e. a *perfect map*.

Underlying reasoning. There are two underlying types of reasoning that are relevant in building Bayes nets: *deductive* and *abductive* reasoning (Winston, 1984; Charniak and

McDermott, 1985). Deductive reasoning is reasoning from causes to effects. Abductive reasoning moves from effects to causes. This is illustrated in Figure 4. Observing rain and predicting that the streets will be wet, is an example of deductive reasoning. The abductive reasoning analog would be observing wet streets and inferring that it must have rained. In a spontaneous interview situation, people often use the same causal syntax for both deductive and abductive statements, which means that both representations of our examples can be reasonably expected to occur in a raw causal map. Therefore, the linkages in the raw causal maps had to be audited and the true direction of causality was established in cooperation with the expert. The direction in causal Bayes nets should reflect the underlying causality, rather than the language used. This is particularly challenging in the case of latent, unobservable concepts, such as 'management quality' where reflective and formative indicators are often not easily distinguishable (Hulland, 1999).

Insert Figure 4 about here

Direct and Indirect Relationships. Standard methods for deriving causal maps do not provide for a distinction between 'direct' and 'indirect' linkages between concepts (Eden et al., 1992; Laukkanen, 1994). In our experience, people, in conversation, often draw direct linkages, even if they are aware of important mediators, since they want to emphasize the fact *that* two concepts are linked and not necessarily *how* they are linked. Thus, for a given pair of variables both direct and indirect connections often appear in interview transcripts (see Figure 5). However, as discussed above, since a Bayesian network is an I-map, the presence or absence of the direct link between two concepts does have implications in terms of conditional independence. If the relationship between "involvement of strong partner" and "decision to invest" is fully mediated by "confidence in timing", the direct arrow needs to be removed. Consequently, the exact nature of such relationships had to be discussed and verified with the domain expert since conditional independencies are critical in making inferences in Bayes nets.

Insert Figure 5 about here

Eliminating circular relationships. The final step in the preprocessing of the raw causal map concerns circular relationships. Causal maps, while directed graphs, have no restrictions with respect to circular relationships. In contrast, Bayesian networks are hierarchical graphs that are by necessity acyclic. Thus, circular relationships need to be removed from the raw causal map. In our raw causal map circular relationships were not present.

At the end of this procedure, the pre-processed map was once again audited by the research participant before probabilities were elicited.

Assignment of states to variables. Each concept from the raw causal map was assigned two states, such as High/Low, True/False, etc. Due to the complex calculations involved in probabilistic inference, such crude state spaces are often unavoidable (see also, e.g., Shepherd & Zacharakis, 2002 for a similar example using conjoint analysis). In addition, Bayesian networks allow probabilities to be assigned to each variable state. Thus, once, management quality, for example, is observed, it can be set to high, low, or to, say, high-low = 80 % - 20 %, allowing for finely graded judgements by the model user.

Assignment of probabilities to states. While the procedures outlined above established the structure of the Bayesian network, the inference processed is driven by a set of prior (for nodes without parents) and conditional probabilities (for nodes with at least one parent) that are attached to each state of each variable. As shown in Figure 2, these are contained in tables. Thus, absent other information, the probability of the management quality being high is 25%. Similarly, the probability that an investment is made given high revenue potential and high quality of management, is 90 %.

We elicited these probabilities directly from the expert. This is a lengthy and complex task. To facilitate the process and to ensure data quality we employed several techniques. One, we used verbal and experiential anchors (shown in Table 1) since most people do not routinely think in probabilities. Second, we focused on one or two sub-sections of the map at a time. Third, we attempted to exploit noisy-AND or noisy-OR (Henrion, 1989) relationships. This is similar to the ability to reduce the number of attribute combinations that need to be directly assessed in conjoint analysis, given assumptions about orthogonality and interaction effects. However, in our example such modeling was usually inappropriate given the complex and asymmetric inter-dependencies between variables. Fourth, we continuously checked for unusual results and had them verified by the venture capitalist. Overall, the elicitation process is iterative in nature. The resulting Bayesian network is depicted in Figure 6 and was implemented using Netica (Norsys Software Corp., 1998), a widely used and user friendly Bayesian Network software package.

Insert Figure 6 about here

Refinement and validation. In addition to general audits, we used real world cases to verify the face validity of the model. We asked the research participant to think of and describe three actual investment opportunities from the past, one successful, one unsuccessful, and one were he decided not to invest. For each variable for which information was available at the time of the decision, we entered his original assessment of the state of those variables into the model. If correct, the probability assigned for the state ‘true’ for the variable ‘Invest’ should reflect the actual decision and the decision confidence of the research participant. This was indeed the case. Figure 7 shows the probabilities for one of these cases (dark gray boxes denote observed / assessed variables). As can be seen, the investment probability was 71.4, supporting the original investment decision. As an additional check we re-entered the states of the variables, this time, however, not the state as assessed at the time of decision, but the actual/real state as revealed in hindsight. To the extent that the decision maker is indeed a competent venture capitalist, the investment probability given by the model now should correspond to the economic fate of the venture. This was the case. In a case from the energy sector, the investment probability based on the information at the time of decision was 57.4, supporting the investment decision. However, merely entering the true information regarding the market potential, reduced the investment probability to 25.4, in line with the disastrous outcome of the actual venture. The validation process also lead to minor refinements to the model. Throughout the process, the VC commented several times (without prompting) on the sometimes surprising, but after reflection appropriate nature of the Bayesian causal map, further increasing our confidence in the face validity of the model.

Insert Figure 7 about here

APPLICATIONS

We propose causal Bayesian networks as a useful tool to complement other bootstrapping decision model methods and decision aids. It can be used multiple ways. First, it can be used to support initial decision making. Evidence from prior studies suggests that in over 90 % of cases, decision makers are outperformed by their own bootstrapping models due to the elimination of unsystematic errors (cf. e.g., Shepherd & Zacharakis, 2002). However, decision aids are just that, giving an initial assessment that serves as a point of departure, rather than replacing decision makers. In addition, immense skill is still required in assessing the states of the variables, many of which are intangible and/or difficult to observe. However, a Bayesian network makes transparent the drivers behind the overall assessment and the software implementation allows easy what-if analysis by changing variable states and observing the automatically updated investment probability.

Second, the mere act of explicating and formalizing hitherto tacit decision models surfaces hidden assumptions, which can now be scrutinized. Research on causal mapping suggests, that the act of drawing a causal map can reduce decision bias (e.g., Hodgkinson et al., 1999). Our research participant also repeatedly commented (without prompting) on the high levels of perceived learning and benefit from the elicitation process itself, before the model was ready for actual use.

Third, novices become experts over time by learning from experience. Bayesian causal maps support this process by making variable assessments and decision drivers explicit and storing them so that they can be compared to reality later. Entering actual rather than predicted values for variables and comparing them, as well as the resulting investment probabilities, can uncover both errors in variable state assessments, and errors in the decision model which can then be refined.

Finally, Bayesian causal maps can be used as points of reference and departure in coaching, teaching, training, and in collaborative learning contexts. Bayesian causal maps can also improve the collaboration of partners within a venture capital firm by enabling them to understand their differences in terms of decision models, biases, and focus. Thus, Bayesian causal maps can help to constructively resolve situations in which partners consistently differ in their assessment of an investment opportunity.

ADVANTAGES OF BAYESIAN CAUSAL MAPS

The use of Bayesian causal maps as decision aids has several advantages. First, in comparison with the hypothetical cases or pure attribute lists employed in many other techniques, the elicitation is based on *real* cases that are meaningful to the respondent, improving the confidence in the meaningfulness of the response. Second, the elicitation of the model structure and the initial pool of variables through causal mapping of multiple case narratives is relatively unobtrusive and less susceptible to bias than direct questions such as: “What factors do you consider when making an investment decision?” Third, with the exception of circular relationships, the method imposes no a-priori assumptions about orthogonality of variables or the (non)-existence of interaction terms. As discussed, any independence assumptions in the final model can be read from the structure of the model by visual inspection. In addition, no symmetric or linear relationships are assumed. For example, in Figure 2, the impact of the observed state of Management Market Know-how on Management Quality depends on whether the observed Market Know-how is high or low. The relationship is asymmetric. Thus, complex interdependencies can be modeled. We observed such asymmetries throughout the model.

Fourth, Bayesian networks can accommodate partial information and uncertainty. Full information is not required. The user can enter his or her assessment for only the variables for which information is known. In addition, assessments can be entered in probabilistic rather than absolute terms (e.g., management quality high / low = 60% / 40%). Fifth, the model is dynamic in the sense that it can be easily updated with new information and the impact of such information is automatically propagated through the model. Sixth, a factor that should not be underestimated is the fact that software is readily available to support the creation - and more importantly – the use of the model. Features that support visualization of, and easy mouse-click interaction with, the model are key to user friendliness and actual use of such models as decision aids in real-world settings.

LIMITATIONS & CHALLENGES

In addition to the limitations highlighted in our discussion of potential model applications, there are several additional limitations and challenges inherent in the construction and use of Bayesian causal maps for venture capital decision making. First, if done conscientiously, the iterative elicitation process can be quite time consuming. This is potentially problematic, given the time constraints faced by venture capitalists.

Second, depending on the structure of the network, the size of the conditional probability tables may become unmanageable. For a node with 8 parents with 2 states each, 256 conditional probabilities would have to be assessed. However, this problem similarly applies to conjoint analysis and other methods. Given certain assumptions, methods such as noisy-AND/OR do exist to reduce the number of probabilities that have to be elicited (e.g., Henrion, 1989).

Third, in this paper we have chosen to elicit conditional probabilities directly. Prior research (Zacharakis & Meyer, 1998) shows that while VCs understanding of their own decision processes is overall quite accurate, it is clearly imperfect.

Fourth, the fact that the original model is based on actual cases is in general a benefit. However, it may also cause idiosyncratic learning from these cases to skew the model. For example, issues that are peripheral in every-day decision making (e.g., the technological and market progress of big industry players), may feature prominently in the model, since it featured prominently in one of the real-life cases used to construct the network. While this may lead to the inclusion of less than salient factors in the model, it does not in itself represent a significant problem.

Fifth, the universe of variables and interdependencies that could be considered is almost limitless. Any model includes only key variables and dependencies and excludes many more. Together with the crude state space (discussed earlier), this may make acceptance of the model by the decision maker more difficult. However, the probabilistic uncertainty that is included in Bayesian causal maps explicitly accounts for such exogenous factors, albeit imperfectly.

Sixth, we believe that the final probability for the "invest" variable, calculated by the model, cannot and should not be interpreted as a probability in the classical sense. For one thing, because of the uncertainty included in the conditional probability tables to account for exogenous factors, there is a floor and a ceiling on the probability of investment inherent in each model. For example, in our model these were 2.91 and 81.4 percent respectively.[2] There are several approaches to this problem. One, one can normalize the final investment probability over the range of possible probabilities: $P_{normalized} = (P_{invest} - P_{floor})/(P_{ceiling} - P_{floor})$. Second, using real cases and past experience, one can construct a verbal translation of probability ranges (e.g., 60%-70% = invest only after serious due diligence). Third, since this model is built to support the decision making of one individual, that venture capitalist will develop his or her own intuition about the meaning attached to the investment probability after using the model for past cases and future projects.

Seventh, our model explicitly models only the salient decision model of a particular venture capitalist. That is, any systematic imperfection in his or her expertise is included in the

[2] Ceilings (floors) were calculated by setting the management quality, product potential, market potential, strong partnerships, and dilution risk variables to their most beneficial (detrimental) state.

model. Also, the model is not transferable across venture capitalists, venture capital firms, or stage of investing.

In the end, decision aids live and die by their acceptance by the venture capitalists that will use them. Given resistance to objectively superior models in the medical field and the fact that the average business plan initially only receives 8-12 minutes attention (Sandberg, 1986), pragmatic as well social and cognitive legitimacy issues may well be the biggest obstacle to the application of the present research.

FUTURE RESEARCH AND CONCLUSIONS

There are several ways the research presented here can be extended. First, the model can be refined to accommodate a wider variety of variables and interdependencies. Second, the modeling exercise could be repeated with a number of participants and systematic differences (across VC firms, novice vs. experts, stage of investment, etc.) could be compared to arrive at a descriptive model of VC investing. Third, it would be interesting to compare the performance of expert knowledge based Bayesian causal maps to pure statistical models such as neural networks. Fourth, we believe that combining causal mapping to elicit the network structure with conjoint or regression techniques to elicit the conditional probability relationships would be a very fruitful avenue for future research. Finally, incorporation of insights from Spohn's theory of epistemic beliefs (Spohn, 1988) would allow a decision model based on order of magnitude probabilities which may be easier to elicit than classical probabilities.

In this paper we have introduced Bayesian causal map as a decision aid in VC decision making. Together with Shepherd and Zacharakis (2002) we are one of the first to move to a critical new step in venture capital decision research and to focus primarily on decision aids and improving decisions rather than focusing on descriptive models. This paper also represents the first implementation and application of Bayesian causal maps first proposed by Nadkarni and Shenoy (2001), thus contributing to the causal mapping, artificial intelligence, and decision modeling literatures. We believe that the study of decision aids in general, and of Bayesian networks in particular, holds great promise for venture capital research.

REFERENCES

Axelrod, R. 1976. *Structure of Decision: The Cognitive Maps of Political Elites*. Princeton, NJ: Princeton University Press.

Banxia Software Limited. 2000. *Decision Explorer 3.1*. Glasgow, UK.

Baron, R. A. 1998. Cognitive mechanisms in entrepreneurship: Why, and when entrepreneurs think differently than other people. *Journal of Business Venturing*, 13: 275–294.

Barr, P. S. 1998. Adapting to unfamiliar environmental events: A look at the evolution of interpretation and its role in strategic change. *Organization Science*, 11: 644–669.

Barr, P. S., & Huff, A. S. 1997. Seeing isn't believing: Understanding diversity in the timing of strategic response. *Journal of Management Studies*, 34: 337–370.

Barr, P. S., Stimpert, J. L., & Huff, A. S. 1992. Cognitive change, strategic action, and organizational renewal. *Strategic Management Journal*, 13: 15–36.

Bonham, G. M., & Shapiro, M. 1976. Explanation of the unexpected: The Syrian intervention in Jordan in 1970. In R. Axelrod (Ed.), *Structure of Decision: The Cognitive Maps of Political Elites*: 113-141. Princeton, NJ: Princeton University Press.

Bougon, M. K., Weick, K. E., & Binkhorst, D. 1977. Cognition in organizations: An analysis of the Utrecht Jazz Orchestra. *Administrative Science Quarterly*, 22: 606–639.

Brunswik, E. 1955. Representative design and probabilistic theory in a functional psychology. *Psychological Review*, 62: 193-217.

Brunswik, E. 1956. *Perception and the representative design of experiments*. Berkley, CA: University of California Press.

Calori, R., Johnson, G., & Sarnin, P. 1994. CEOs' cognitive maps and the scope of the organization. *Strategic Management Journal*, 15: 437–457.

Carley, K., & Palmquist, M. 1992. Extracting, representing and analyzing mental models. *Social Forces*, 703: 601–636.

Charniak, E., & McDermott, D. 1985. *Introduction to Artificial Intelligence*. Reading, MA: Addison-Wesley.

Cyert, R.M., & March, J.G. 1963. *A Behavioral Theory of the Firm*. Englewood Cliffs, NJ: Prentice Hall.

Eden, C. 1991. Working on problems using cognitive mapping. In S. C. Littlechild & M. Shutler (Eds.), *Operations Research Management*: 236–259. London: Prentice-Hall.

Eden, C., & Ackermann, F. 1993. Evaluating strategy—Its role within the context of strategic control. *Journal of Operational Research Society*, 44: 853–865.

Eden, C., & Ackermann, F. 1998. *Making Strategy: The Journey of Strategic Management*. London: Sage.

Eden, C., Ackermann, F., & Cropper, S. 1992. The analysis of cause maps. *Journal of Management Studies*, 293: 309–323.

Eden, C., Jones, S., & Sims, D. 1983. *Messing about in Problems*. Oxford: Pergamon.

Eden, C., Jones, S., Sims, D., & Smithin, T. 1981. The intersubjectivity of issues and issues of intersubjectivity. *Journal of Management Studies*, 18: 37–47.

Eden, C., & Spender, J.-C. (Eds.). 1998. *Managerial and Organizational Cognition: Theory, Methods and Research*. Thousand Oaks, CA: Sage.

Fahey, L., & Narayanan, V. K. 1989. Linking changes in revealed causal maps and environmental change: An empirical study. *Journal of Management Studies*, 26: 361–378.

Fiol, C. M. 1989. A semiotic analysis of corporate language: Organizational boundaries and joint venturing. *Administrative Science Quarterly*, 34: 277–303.

Fiol, C. M. 1990. Explaining strategic alliance in the chemical industry. In A. S. Huff (Ed.), *Mapping Strategic Thought*: 227–249. Chichester: Wiley.

Fiol, C. M. 1995. Corporate communications: Comparing executives' private and public statements. *Academy of Management Journal*, 38: 522–536.

Fiol, M., & Huff, A. S. 1992. Maps for managers: Where are we? Where do we go from here? *Journal of Management Studies*, 29: 269–285.

Hall, J., & Hofer, C. W. 1993. Venture Capitalists' Decision Criteria in New Venture Evaluation. *Journal of Business Venturing* 8: 25–42.

Hall, R. I., & Menzies, With B. 1983. A corporate system model of a sports club: Using simulation as an aid to policy making in a crisis. *Management Science*, 29: 52-64.

Henrion, M. 1989. Some practical issues in constructing belief networks. In L. N. Kanal, T. S. Levitt, & J. F. Lemmer (Eds.), *Uncertainty in Artificial Intelligence*, 3: 161-173. Amsterdam: North Holland.

Hisrich, R. D., & Jankowicz, A. D. 1990. Intuition in venture capital decisions: An exploratory study using a new technique. *Journal of Business Venturing*, 5: 49-62.

Hodgkinson, G. P., Brown, N. J., Maule, A. J., Glaister, K. W., & Pearman, A. D. 1999. Breaking the frame: An analysis of strategic cognition and decision making under uncertainty. *Strategic Management Journal*, 20: 977–985.

Hoffman, P. J. 1960. The paramorphic representation of clinical judgment. *Psychological Bulletin*, 57: 116-131.

Huff, A. S. (Ed.). 1990. *Mapping Strategic Thought*. Chichester: Wiley.

Hulland, J. 1999. Use of partial least squares (PLS) in strategic management research: A review of four recent studies. *Strategic Management Journal*, 2: 195-204.

Langfield-Smith, K., & Lewis, G. 1989. *Mapping cognitive structures: A pilot study to develop a research method*. Working Paper No. 14, The University of Melbourne Graduate School of Management, Melbourne.

Laukkanen, M. 1994. Comparative causal mapping of organizational cognition. *Organization Science*, 53: 322–343.

Lauritzen, S. L., Dawid, A. P., Larsen, B. N., & Leimer, H.-G. 1990. Independence properties of directed Markov fields. *Networks*, 205: 491–505.

MacMillan, I. C., Zemann, L., & Narasimha, P. N. S. 1987. Criteria distinguishing unsuccessful ventures in the venture screening proposals. *Journal of Business Venturing*, 2:123–137.

MacMillan, I. C., Siegel, R., & Narasimha, P. N. S. 1985. Criteria used by venture capitalist to evaluate new venture proposals. *Journal of Business Venturing*, 1: 119–128.

Muzyka, D., Birley, S., & Leleux, B. 1996. Trade-offs in the investment decisions of European venture capitalists. *Journal of Business Venturing*, 11: 273-287.

Nadkarni, S., & Shenoy, P. P. 2001. A Bayesian network approach to making inferences in causal maps. *European Journal of Operational Research*, 128: 479–498.

Naylor, J. C., & Wherry, R. J., Sr. 1965. The use of simulated stimuli and the “JAN” technique to capture and cluster the policies of raters. *Educational and Psychological Measurement*, 25(4): 969-986.

Nelson, K., Nadkarni, S., Narayanan, & V. K., Ghods, M. 2000. Understanding software operations support expertise: A revealed causal mapping approach. *MIS Quarterly*, 24: 475–507.

Newell, A., & Simon, H. 1972. Human Problem Solving. Englewood Cliffs, NJ: Prentice-Hall.

Norsys Software Corp. 1998. *Netica for Windows, 1.12*. Vancouver, BC, Canada.

Pearl, J. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, CA: Morgan Kaufmann.

Riequelme, H., & Rickards, T. 1992. Hybrid conjoint analysis: An estimation probe in new venture decisions. *Journal of Business Venturing*, 7: 505-518.

Ruhnka, J. C., Feldman, H. D., & Dean, T. J. 992. The "living dead" phenomena in venture capital investments. *Journal of Business Venturing*, 7(2): 137-155.

Russell, R. D. 1999. Developing a process model of intrapreneurial systems: A cognitive mapping approach. *Entrepreneurship: Theory & Practice*, 23(3): 65-84.

Sandberg, W. R. 1986. New venture performance: The role of strategy, industry structure, and the entrepreneur. Lexington, MA: Lexington Books.

Sandberg, W. R., Schweiger, D. M., & Hofer, C. W.1988. The use of verbal protocols in determining venture capitalist's decision processes. *Entrepreneurship Theory & Practice*, 13(2): 8-20.

Schwenk, C. R. 1984a. Cognitive Simplification Processes in Strategic Decision-Making. *Strategic Management Journal*, 5: 111-128.

Schwenk, C. R. 1984b. Effects of planning aids and presentation media on performance and affective responses in strategic decision-making. *Management Science*, 30: 263-272.

Shepherd, D. A. 1999. Venture capitalists' assessment of new venture survival. *Management Science*, 455: 621–632.

Shepherd, D. A., & Zacharakis, A. 2002. Venture capitalists' expertise: A call for research into decision aids and cognitive feedback. *Journal of Business Venturing*, 17: 1-20.

Simon, H. A. 1955. A Behavioral Model of Rational Choice. *Quarterly Journal of Economics*, 16(1): 99-118.

Spiegelhalter, D. J., Dawid, A. P., Lauritzen, S. L., & Cowell, R. G. 1993. Bayesian analysis in expert systems. *Statistical Science*, 83: 219–283.

Spohn, W. 1988. Ordinal conditional functions: A dynamic theory of epistemic states. In W. L. Harper & B. Skyrms (Eds.), *Causation in Decision, Belief Change, and Statistics*, 2: 105-134. Dordrecht, Holland: D. Reidel.

Stewart, T. R. 1988. Judgement analysis: Procedures. In B. Brehmer & C. Joyce (Eds.), *Human Judgement: The SJT View*. Amsterdam: North Holland.

Tolman, E. C. 1948. Cognitive maps in rats and men. *Psychological Review*, 55: 189–208.

Tyebjee, T. T., & Bruno, A. V. 1984. A model of venture capitalist investment activity. *Management Science*, 30: 1051–1066.

Winston, P. H. 1984. *Artificial Intelligence*. Reading, MA: Addison-Wesley.

Wrightson, M. T. 1976. The documentary coding method. In R. Axelrod (Ed.), *Structure of Decision: The Cognitive Maps of Political Elites*: 291–332. Princeton, NJ: Princeton University Press.

Zacharakis, A. L., & Meyer, G. D. 1995. The venture capitalist decision: Understanding process versus outcome. *Frontiers of Entrepreneurship Research*, 465-478.

Zacharakis, A. L., & Meyer, G. D. 1998. A lack of insight: Do venture capitalists really understand their own decision process? *Journal of Business Venturing*, 13: 57–76.

Zacharakis, A. L., & Meyer, G. D. 2000. The potential of actuarial decision models: Can they improve the venture capital investment decision? *Journal of Business Venturing*, 154: 323-346.

Zacharakis, A. L., & Shepherd, D. A. 2001. The nature of information and overconfidence on venture capitalists' decision making. *Journal of Business Venturing*, 16: 311–332.

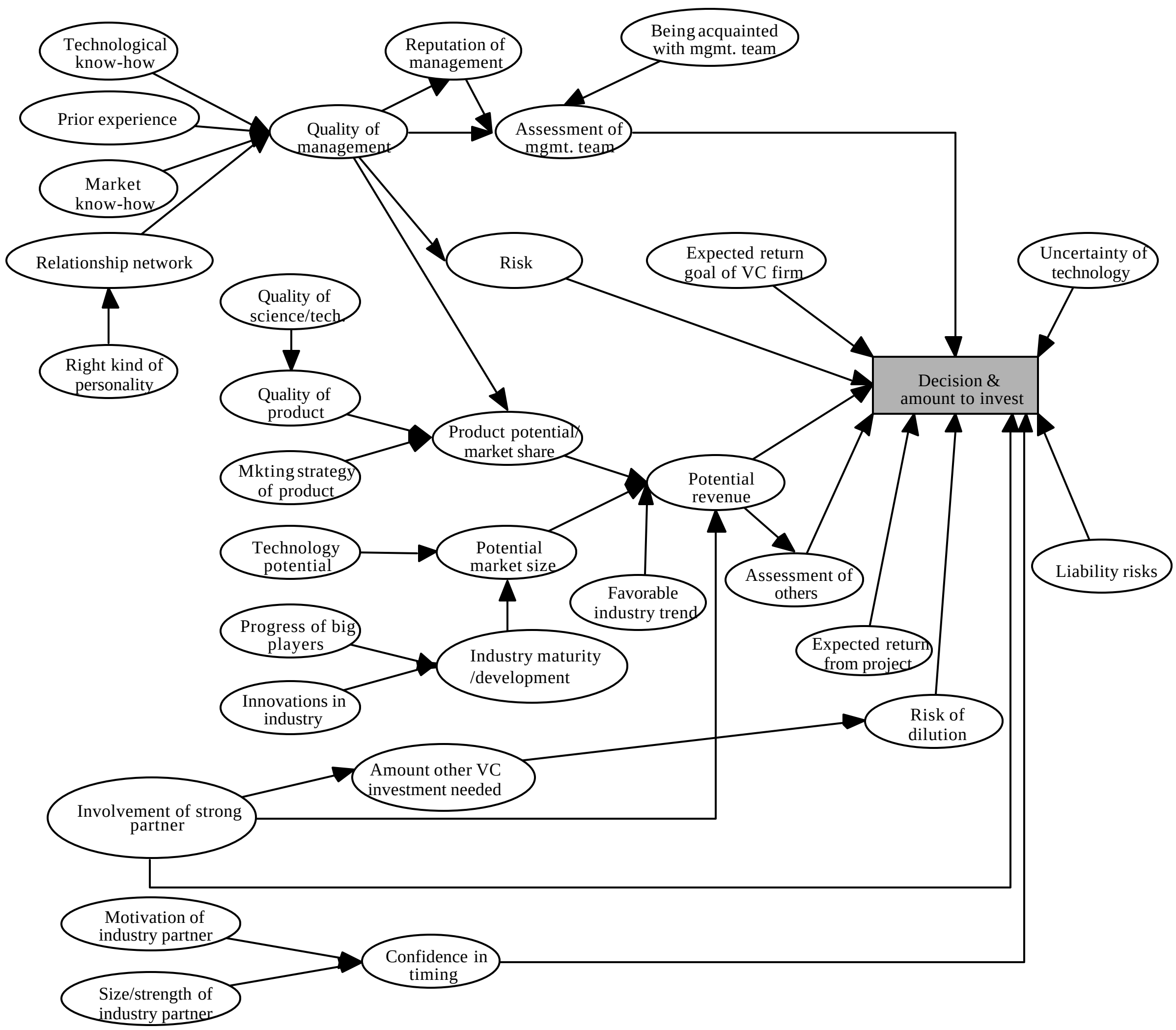
TABLE 1:

Verbal and Experiential Probability Anchors

Verbal Descriptor	Experiential (Coin-Toss) Anchor	Probability
very likely	1 or more heads on 3 tosses	7/8
likely	1 or more heads on 2 tosses	3/4
even	heads on a toss	1/2
unlikely	2 heads on 2 tosses	1/4
very unlikely	3 heads on 3 tosses	1/8

FIGURE 1:

Raw Causal Map‡



‡ **Note**: Plus and minus signs above arrows have been omitted to improve readability; the directionality of the relationship is mostly self-evident.

FIGURE 2:

A Bayesian Network with Conditional Probability Tables

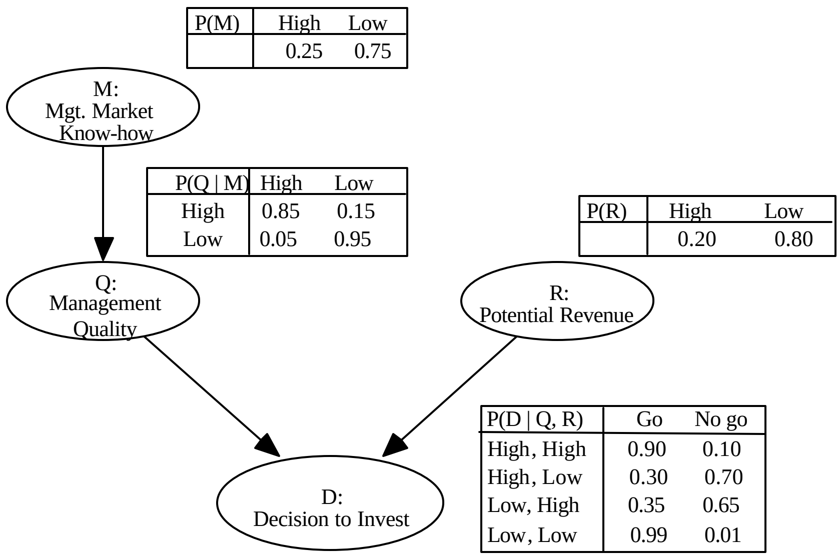


FIGURE 3:

A Positive Causal Relation Between Two Variables

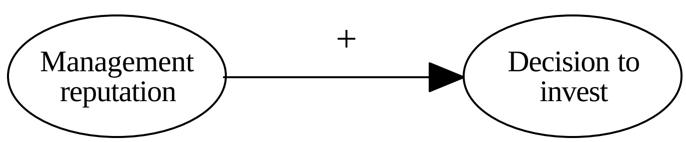


FIGURE 4:

Examples of Deductive and Abductive Reasoning

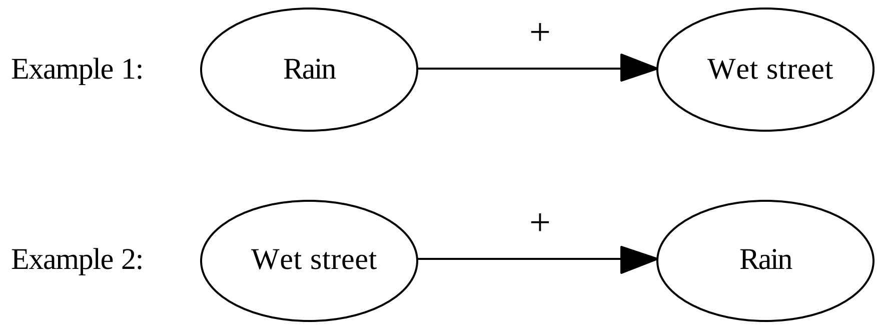


FIGURE 5:

An Example of Direct and Indirect Reasoning

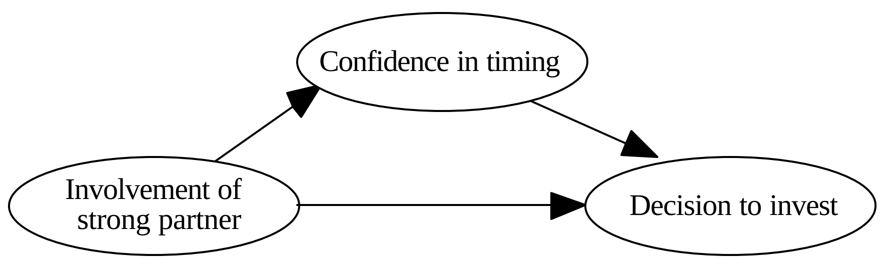


FIGURE 6:

Bayesian network - Base case of no information

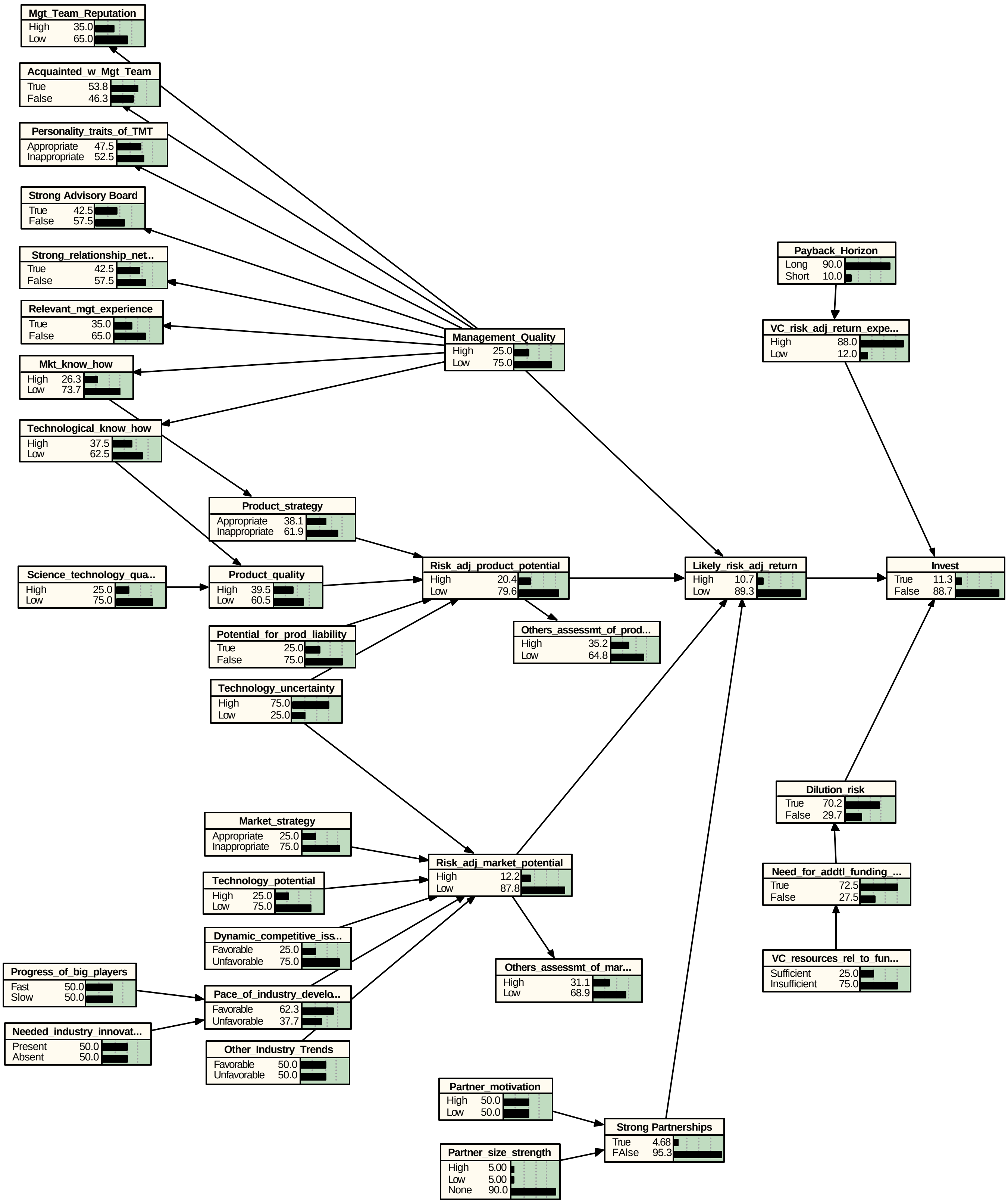


FIGURE 7:

Bayesian network - Chemical industry case