Towards the Accuracy of Cybernetic Strategy Planning Models: Causal Proof and Function Approximation

Christian A. HILLBRAND
University of Applied Sciences Liechtenstein
LI-9490 Vaduz, Principality of Liechtenstein

ABSTRACT

All kind of strategic tasks within an enterprise require a deep understanding of its critical key success factors and their interrelations as well as an in-depth analysis of relevant environmental influences. Due to the openness of the underlying system, there seems to be an indefinite number of unknown variables influencing strategic goals. Cybernetic or systemic planning techniques try to overcome this intricacy by modeling the most important cause-and-effect relations within such a system. Although it seems to be obvious that there are specific influences between business variables, it is mostly impossible to identify the functional dependencies underlying such relations. Hence simulation or evaluation techniques based on such hypothetically assumed models deliver inaccurate results or fail completely.

This paper addresses the need for accurate strategy planning models and proposes an approach to prove their cause-and-effect relations by empirical evidence. Based on this foundation an approach for the approximation of the underlying cause-and-effect function by the means of Artificial Neural Networks is developed.

Keywords: Strategy planning, universal function approximation, causal proof, Artificial Neural Networks, Balanced Scorecard, cause-and-effect relations, cybernetic knowledge, Sensitivity Model.

1. INTRODUCTION

Decision-makers working on a corporate planning issue face one tremendous dilemma: In order to determine an optimal direction for the future strategy of an enterprise or even to deduce correct measures, they need to thoroughly analyze the questions underlying this problem. On the other hand, for most corporate decisions it is necessary that they can be settled within reasonable time which creates a certain clash of interests: Very often managers are not able to decide optimally on specific issues because they are lacking analytical possibilities within this limited time-span. As it is evident, there are hardly any potential measures to analyze all influences and consequences of a strategic or tactic decision in a daily business routine. The reason for this inability to fully explain the environment of such an issue is the massive complexity of interwoven factors corresponding with a certain corporate measure.

Several approaches try to reduce this complexity by modeling the underlying cause-and-effect relations between certain business elements.

A well known strategic planning methodology based on these principles – the Balanced Scorecard – was introduced by KAPLAN and NORTON [18]: In order to formulate a homogenous business strategy they propose to harmonize individual goals by linking them together within cause-and-effect diagrams. The logic behind this technique is based on their findings that key business measures can most efficiently be influenced via so-called lagging indicators or business drivers [19], which are mostly of non-financial nature. As a consequence, they postulate to introduce new measures out of four distinct perspectives in order to build a holistic strategic management framework. According to KAPLAN and NORTON, managers should formulate strategic goals which are equally distributed over all of the perspectives – financial, as well as process-, customer- or innovation-specific. The cause-and-effect relations – linking these measures together – allow the identification of competing goals:

> *"By forcing senior managers to consider all the important operational measures together, the balanced scorecard lets them see whether improvement in one area may have been achieved at the expense of another." [18]*

Another approach is derived from the work of VESTER (see [32], [33] or [34]) in the area of biocybernetic complexity management: His Sensitivity Model depicts a cybernetic system with dynamic influence relations between interacting elements. This basic concept is summarized as follows:

> *"One of the main reasons for the crisis of our industrial society lies in the lack of awareness of the closely interwoven factors, which are involved in the process of our civilization." [34]*

The causal relations – as a part of the Sensitivity Model – can be easily transformed into a cross-impact-matrix, which contains influence indices between pairs of elements. As a consequence it is possible to derive two important characteristic attributes of an element: The Active Sum (AS) is determined by summing up all indices within a row and can be interpreted as the overall effect, a specific element has on the whole system. On the contrary the sum of the concerning table column is known as the Passive Sum (PS) of this element: It is a measure of the impact of the system on a specific variable. Comparing the magnitude of the Active vs. the Passive Sum, one can make statements about the activity of an indicator within the cybernetic system. VESTER [34] classifies elements with the property AS>PS as active element. Correspondingly measures are called reactive where AS < PS.

When applied to strategic planning issues, managers are frequently facing the challenge to find strategic goals, which have the highest possible impact on the overall system. The ratio between necessary expenses for the achievement of a certain objective and the indirect effects associated with them can be optimized consequently.

The two approaches, which have been illustrated so far, provide powerful tools to tackle the complexity, which is characteristic for strategic planning as an optimization technique applied to a cybernetic system. However, they have one crucial shortcoming in common since they are based on the assumption that the cause-and-effect relations are given. As a consequence they do not describe a way to ascertain objective influence measures:

"Over the short term, managers' assessment of strategic impact may have to rest on subjective and qualitative judgments. Eventually, however, as more evidence accumulates, organizations may be able to provide more objectively grounded estimates of cause-and-effect relationships." [19]

As the proof of existence and the quantitative characteristics of these dependencies are vital for the quality of the overall model and the strategic decisions based on it, this seems to be a substantial source of errors.

In order to overcome these shortcomings of traditional causal models in strategic planning, this paper proposes an approach which provides techniques to empirically support the accuracy of cause-and-effect relations: A fundamental analysis of the corporation's "cybernetic knowledge" – it's key figures – can provide powerful insights into the system of interlinked variables, their correlation and causation. The definition of a modeling framework for causal strategy models in section 2 represents the foundation for the discussion of techniques to identify correlation and causation in graphical models based on empirical data (section 3). An overall approach to evaluate hypothetic cause-and-effect relations within a strategy model by the means of Artificial Neural Networks is proposed in section 5.

2. A MODELING FRAMEWORK FOR CAUSAL STRATEGY MODELS

Before discussing details about the inference of quantitative statements in a causal strategy model it is necessary to provide the description of a stable foundation for modeling cybernetic systems: The purpose – discussed so far – requires syntactic rules for the definition of a cause-and-effect model in the form of a meta-model.

As a prerequisite for the formulation of the modeling framework it is necessary to describe a consistent understanding of a holistic business ratio system: Besides quantitative methods which are commonly known from traditional ratio systems (mostly in financial accounting) it seems to be necessary to introduce alternative techniques in order to take the effects of lagging indicators (or active elements) into account. In the terms of the Balanced Scorecard methodology these variables – responsible for sustained growth of financial measures – very often characterize the quality of the internal business processes, as well as customer- or innovation-related topics respectively. Hence they are very often of a fuzzy nature and – as a consequence – cannot be measured by traditional means.

Therefore one can identify the following key elements of a holistic business ratio system underlying the modeling framework: Indicators (see Definition 1) and influence relations (see Definition 2) connecting them.

> **Definition 1:** An **indicator**, **variable**, **measure** or **ratio** is an instrument to measure the status of a certain strategic, operational or tactic (sub)goal of an organization. Every indicator is related to a frame of reference, which allows making qualitative statements about how to interpret this measure regarding the underlying goal.

An indicator i is of crisp nature, if it can be measured definitely by the means of a scale of measurement, as it is the case for all traditional financial measures. Following the taxonomy of STEVENS [29], which is most common in statistical literature (see [2] and [6]), one can identify four different measurement scales:

- **Nominal**, also known as **qualitative** or **categorical variables**, are discrete attributes by which the characterized unit can be classified but not ranked or ordered (e.g.: yellow, green, blue).
- **Ordinal** or **ranked variables** are discrete attributes by which the characterized unit can be ordered using an ordinal scale (e.g.: small < medium < large).
- **Interval variables** have ordinal properties and can be measured on a continuous scale, which allows deriving meaningful differences between two elements. For example a person with an IQ of 70 is not necessarily half as intelligent as an individual characterized by an IQ of 140 (following [2]).
- For **ratio variables**, the additional property that the ratio of two indicators is meaningful holds. This category covers most of the indicators encompassed by classical models: Earnings of $ 1 billion are twice as much as $ 500 million.

Because the value of a nominal or ordinal variable cannot be assessed exactly, henceforth the term fuzzy indicator ($\tilde{\imath}$) will be used for the rest of this paper. In contrast crisp measures (i) are characterized by their ability to be measured on a continuous numerical scale and hence are covered by the classes of interval or ratio variables respectively. Most traditional decision supporting techniques are focused on these crisp ratios. As a consequence the lack of numeric measurability of fuzzy variables leads to the need for special operations dealing with this type of indicators. Alternatively they have to be mapped to a corresponding crisp indicator $\tilde{\imath} \rightarrow i^{corr}$ (e.g. express customer satisfaction as an index or percentage).

Since it is not sufficient to know the values of an organization's indicators for making strategic decisions, the knowledge about the interaction between those ratios is crucial. As a consequence the influence on an indicator can be defined as follows:

> **Definition 2:** The **influence** $\varphi(I^{in}, i^{res})$ of a set of incoming measures I^{in} on a result ratio i^{res} indicates the proportion and the degree of sensitivity by which the latter reacts on changes of one or more incoming indicators.

If the influence of a single incoming measure can be isolated, it is defined as the relation or sensitivity between this indicator and the result ratio, which can be described formally as follows: Let $f(I^{in})$ be the effect, the incoming set I^{in} has on the result indicator i^{res}, then

$$\begin{aligned} &\varphi_f\left(i_j^{in}, i^{res}\right) = \Delta i^{res} = \\ &= f\left(I^{in}{}'\right) - f\left(I^{in}\right) = \\ &= f\left(i_1^{in}, \ldots, i_j^{in} + 1, \ldots, i_n^{in}\right) - f\left(i_1^{in}, \ldots, i_j^{in}, \ldots, i_n^{in}\right) \end{aligned} \tag{1}$$

In some cases of influence between sets of crisp indicators it is possible to formally describe a functional dependency by an (in)equation as in the following example:

$$net\ working\ capital = \frac{current\ assets}{short-term\ liabilities} \tag{2}$$

If one defines this type of axiomatically determined association as known or defined influence, all kind of interrelations where either one element of the set of incoming measures or the result indicator is of fuzzy nature are excluded. Furthermore, relations connecting crisp ratios with an unknown degree of influence cannot be described functionally. These two exceptions produce a set of undefined influences which give rise to the postulate for improved evaluation methods.

The transitive properties of associations – as illustrated in eq. (3) – are based on the fact that a specific ratio can be result of an influence relation as well as an element of an incoming indicator set:

$$\begin{aligned} &if && f\left(I^{in_1}\right)=i^{r_1} \\ &and && i^{r_1}\in I^{in_2}: g\left(I^{in_2}\right)=i^{r_2} \\ &then && \exists h\left(I^{in_1}\right)=i^{r_2} \end{aligned} \tag{3}$$

An influence $\varphi_h(I_j^{in1},i^{r2})$ of this type, which is generated by at least two transitive direct influences, is defined as indirect.

The main implication for strategic planning is to use these definitions of influence relations in order to harmonize goals: A consistent strategy should only contain objectives, which are measured by indicators linked by significantly positive reciprocal influences.

As depicted in figure 1, every strategic goal is measured by at least one indicator of crisp or fuzzy type. Following the above categorization of the term "defined influence"; this kind of relation is only valid between two crisp indicators (i.e. interval or ratio variables). It can only be modeled if the underlying functional dependencies are known. Hence there exist axiomatic rules that describe the transformation of one or more influencing indicators into a resulting measure.

Still, there exist cause-and-effect relations where one or both premises for a known influence do not hold. Most decision-supporting systems implementing causal models abandon this type of relation because it cannot be described by the means of a formal expression and therefore it is not possible to evaluate them by traditional algebraic techniques. However, in the proposed modeling framework it is possible to integrate such beliefs in the form of the relation "undefined influence" because they represent an important part of a decision-maker's conception about the causation of key measures: This type of association encompasses most influences of fuzzy non-financial business drivers on lead ratios. Correspondingly the validity of such a relation has to be given between both – "fuzzy" and "crisp indicators". Therefore the abstract meta-model-class "indicator" – inheriting its properties to both types of measures mentioned above – is used as starting and ending point of this reflexive link (see figure 1).

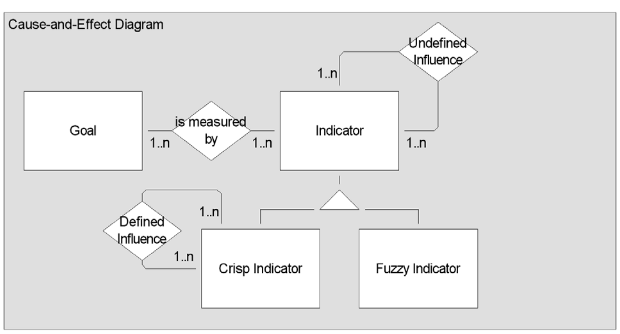

figure 1: Meta-model of a Cause-and-Effect Diagram

The main intention of this paper is to provide some basic prerequisites allowing the evaluation of cause-and-effect relations in strategy models. The two different types of influence associations which have been identified so far require different specific tools to infer assumptions about the underlying rules. Since one of the central characteristics of known relations is its definability in functional terms, the causal laws are evident and can easily be modeled. Therefore the cause-and-effect meta-model is to be extended accordingly: As illustrated in figure 2, the relation "defined influence" is augmented to the state of a class in order to allow refinements of this relation to be described in a sub-diagram called "Indicator Model": Crisp indicators can be joined by the four basic arithmetical operations. The result of this procedure is stored in a "composed indicator" which again can be the input for another operation or serve as an equivalent of a crisp result ratio in the Cause-and-Effect Diagram. Instances of the class "referenced indicator" allow the cross-reference of (composed) variables from other Indicator Models for re-use. The leaves of an expression tree – constructed in this manner – are either of the type "constant", "referenced indicator" or "elementary indicator". The latter can be connected to a data extraction component which models and implements mining paths out of operational databases (for a detailed description of this functionality see [15]).

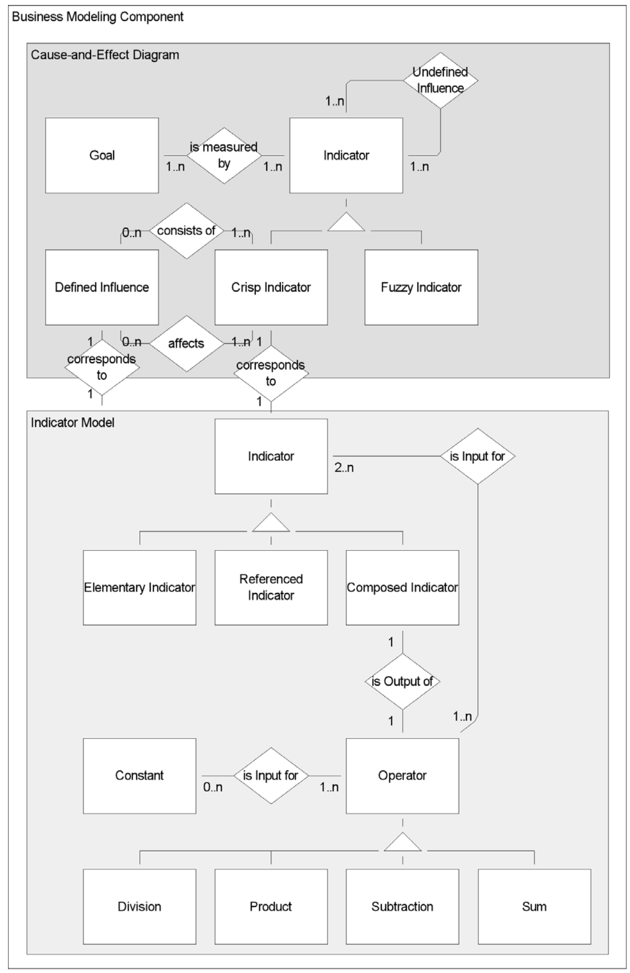

figure 2: Extended meta-model of a Cause-and-Effect Diagram to represent Defined Influences

Although this meta-model allows the description of causal influences with a known functional structure, it seems to be too weak to scrutinize the exact nature of undefined associations as they lack functional definability a priori. However, it is possible to use empirical data to make inferences about these unknown causal relations. There are several approaches known in statistics and Artificial Intelligence that use correlational studies to approximate a rule underlying a hypothetic assumption based on a decision-maker's belief. But mere correlation is only a necessary but not a sufficient prerequisite (conditio sine qua non) for a causal influence between two variables as numerous examples of so-called spurious correlations show. This problem class – which will be addressed in the next section – characterizes associations where no causal dependence can be assumed or where the direction of influence is not clear.

3. IDENTIFYING CORRELATION AND CAUSATION IN GRAPHICAL MODELS

In order to evaluate an undefined causal influence between two or more strategic measures it is necessary to determine whether these parameters correlate and if this relation has causal or non-causal reasons. Associations of the latter type are known as spurious associations dating back to a definition of SIMON [26].
As causality between business variables usually cannot be inferred or observed by axiomatic rules one has to employ a suitable definition in order to empirically prove cause-and-effect relations between sets of indicators. As there does not exist a uniform conception of this term within scientific literature (cf. [13, pp. 48 - 58]) this paper employs four specific conditions for causal relations. By using this concept of causality it is possible to prove variables to be cause and effect of each other by mere empirical observations. Hence it is not necessary to perform costly long-term experiments. These four causality principles are (cf. [13, p. 65]):

- Previous causal knowledge
- Informational redundancy
- Temporal precedence
- Control of third variable effects

By modeling causal assumptions as cause-and-effect hypotheses the requirement of previous causal knowledge is already fulfilled. This prerequisite reduces the set of potentially spurious cause-and-effect relations dramatically compared to a complete system model where all elements are fully interlinked.
The criterion of informational redundancy describes the property of an independent timeseries to contain information which can be used to (partially) explain the variation of a dependent timeseries. In the specific case of linear functional dependencies between two causal variables this property can be measured by the concept of cross correlation. For nonlinear causal functions the performance of this method regarding the ability to identify informational redundancy deteriorates, as experimental results show (cf. [14, pp. 44 - 48]).
As stated before, correlation in specific and informational redundancy in general is a necessary but not sufficient condition for causality. Therefore an important property that distinguishes causal from spurious relations is a temporal sequence between incoming and resulting indicators: PEARL and VERMA state that "temporal precedence is normally assumed essential for defining causation, and it is undoubtedly one of the most important clues that people use to distinguish causal from other types of associations" [23]. As COHEN et al. point out, it is trivial to detect this characteristic in an experimental setting: "If x and y are correlated and y changes if x is manipulated, then x causes y" [3]. However, it becomes a challenging issue to infer causal relations from experience because temporal aspects are not evident from mere correlational studies. Thus, many theories (e.g. [16]; see [23] for a discussion) stated that causal inferences become impossible without background knowledge derived from experiments (i.e. an explicit manipulation of variables).
Disproving these assertions, considerable work has been done on the design of techniques to infer and prove temporal precedence in causal relations only from observation: For example BOX and JENKINS propose their Multivariate AutoRegressive Moving Average approach (MARMA) which constructs a mathematical forecasting model in order to express a given value of a dependent timeseries y_t by its prior values $y_{t-\Delta t}$, by past values of independent variables x_{t-b-i} and by some unknown disturbance value e_t. Applying the concept of cross correlation between pairs of prewhitened lagged timeseries it is possible to identify the model by estimating the linear parameters, as well as the disturbance term, the minimum timelags b and the maximum timelags $b+s$ ($s = max\ i$) of each independent variable.
Extending the definition of correlated and lagged variables by the fourth principle means to rule out the possibility of an unknown third variable influencing both – the independent and the dependent indicator – and therefore inducing a spurious association between the latter two. In order to detect such effects it is necessary to apply more advanced methods to the model graph:
For example PEARL and VERMA [23] propose an algorithm which classifies relations between variables as potential, genuine or spurious causal associations: The IC-Algorithm (Inductive Causation) tests the conditional dependence[1] of a state Z preceding the occurrence of a certain variable Y. If Y is conditionally dependent on Z given another indicator X but independent given a context not including X, it can be assumed that X has a causal influence on Y. Similar approaches have been proposed by IWASAKI and SIMON [17], COOPER and HERSKOVITS [4], as well as SPIRTES et al. [28].

4. THE NEED FOR NONPARAMETRIC METHODS FOR THE UNIVERSAL APPROXIMATION OF CAUSAL FUNCTIONS

As it has been shown in the previous section of this paper it is possible to prove or disprove the causal properties of a given association between a pair of business variables. With this prerequisite one can easily evaluate the cause-and-effect-hypotheses within a cybernetic strategy planning model. But this proof does not provide the required properties for a simulation model. Therefore it is necessary to approximate the causal functions underlying the previously proven cause-and-effect relations.
Statistical methods in this area mostly use analytical approaches to identify some measure of association: The term "covariance" can be regarded as a central concept in this context. It is based on the assumption that two variables X and Y are independent if

$$E(X \cdot Y) = E(X) \cdot E(Y) \tag{4}$$

A violation of this rule leads to the conclusion of dependency between the two indicators. The difference between the two parts of eq. (4) is called covariance and can range from 1 (perfect positive correlation) to –1 (perfect negative correlation):

$$Cov(X,Y) = E(X \cdot Y) - E(X) \cdot E(Y) = \frac{\sum x_i y_i - \overline{xy}}{n} \tag{5}$$

There exist numerous parametric[2] and non-parametric extensions of this principle in the statistical literature, which will not be discussed in this paper (see [12] for details).
Still, those methods seem to have one crucial inadequacy in common: They are relatively sensitive to imperfections within the examined area because most of these techniques require

[1] For a detailed discussion of the conditional independence property see the discussion in [21] and [22].
[2] Parametric Models are characterized by certain assumptions (e.g. normally distributed sample, etc.). Common approaches are the classical forms of Regression Analysis (linear or nonlinear; bi- or multivariate) and the Analysis of Variance (ANOVA).

completely randomized samples as a precondition for their application. In comparatively complex environments such as strategic planning it is scarcely possible to obtain such complete data sets for all measures.

Another obstacle for the application of these concepts for the approximation of causal functions is their restriction to linear dependencies. As numerous observational studies show, it cannot be assumed that all relations between business variables are near to linear. On the contrary one can observe saturation effects as well as scale and similar effects which usually approximate the shape of nonlinear functions. In fact, growth curves very often describe some kind of S-shaped or sigmoidal function as for example BEER observes (cf. [1, pp.10 ff.])

Recapitulating these objections it seems to be crucial for the evaluation of strategic causal functions to employ an approximation algorithm that is fault-tolerant. Furthermore it is not possible to provide a-priori knowledge of any arbitrary function underlying a causal relation as a premise. As this is necessary for a number of methods like regression models, FOURIER or TAYLOR series expansion, it is essential to identify another approach which provides this required property of universal function approximation.

Since the research area of Artificial Intelligence provides some alternative solutions with respect to these issues, this paper follows one specific approach which promises a solution to the disturbance problem as well as the universal approximation postulate: Artificial Neural Networks mimic the functional behavior of neurons which enable a massively parallel processing of information by the brain. Therefore, in the relevant literature they are often characterized as “adaptive and fault tolerating systems for information processing” [20, p. 211].

Furthermore several papers show that Artificial Neural Networks can be employed as universal function approximators. The proof of this property can be traced back to KOLMOGOROV´s superposition theorem [31, p. 187]:

> **Theorem 1 (KOLMOGOROV´s superposition theorem):** For all $n \geq 2$, and for any continuous real function f of n variables on the domain $[0, 1], f : [0, 1]^n \rightarrow R$, there exist $n(2n+1)$ continuous, monotone increasing univariate functions on $[0, 1]$, by which f can be reconstructed according to the following equation:

$$f(x_1,\ldots,x_n) = \sum_{q=0}^{2n} \phi_q \left(\sum_{p=1}^{n} \psi_{pq}(x_p) \right)$$

The bottom line of this theorem is that the combination of (at least) two functions ϕ_q and ψ_{pq} can be used to approximate any arbitrary continuous real function f. Applying this insight, to ANN theory, HECHT-NIELSEN (cf. [8]) shows that KOLMOGOROV´s theorem can be interpreted as a feedforward Multilayer Perceptron (MLP): The transfer function of the hidden layer nodes ψ_{pq} have to be continuous real functions mapping their input to the domain $[0,1]$. Usually these so-called squashing functions are implemented by a sigmoidal or logistic respectively a heaviside activation function (cf. [13, pp. 108 f.]). Incorporating these necessary prerequisites, Artificial Neural Networks can be regarded as universal function approximators which can be applied to any causal business function.

Accordingly, a number of surveys (e.g. [5], [7], [12], [25], [27] or [30]) dealing with a managerial or economic background state that Artificial Neural Networks are characterized by their superior performance over statistical concepts (mostly regression models). DENTON summarizes this fact as follows:

> *“The results of the designed experiment clearly demonstrate the superiority of forecasting with neural networks. Elimination of ambiguities in the independent variables allows the neural network to successfully learn the interaction from the data, without having to specify it in a model definition. Performing a statistical regression with a mis-specified model can result in biased and inconsistent parameter estimates. If the data are unambiguous, the neural network will be unaffected by this flaw”* [5].

For these reasons this paper proposes the application of Artificial Neural Networks for function evaluation in causal strategy models as presented section 5. Dispensing with details concerning the background neural computing, the relevant literature is to be analyzed for further details (e.g. [24]).

5. ESTIMATION OF HYPOTHETIC CAUSE-AND-EFFECT RELATIONS BY THE MEANS OF ARTIFICIAL NEURAL NETWORKS

Section 2 introduces a modeling framework for rudimentary cause-and-effect relations consisting of crisp and fuzzy indicators as well as known and unknown influence relations. A meta-model for the detailed description of axiomatic known influences has been proposed in the form of an Indicator Model. In this section the issue of inferring proven unknown causal functions from empirical observations is to be adressed. Despite the findings of the previous sections, managers seem to take only little effort to substantiate their assumptions about basic cause-and-effect relations. Most Decision or Executive Support Systems only propose to review these suppositions within a team of specialists knowing the internal business structures very well (see [2] for details).

Since this approach is considered as contradictory with regard to its intention – because it tends to over-emphasize the internal perception of temporary difficulties and to neglect external influences on indicators – one has to analyze more deeply the uniqueness of this team-knowledge in order to verify its results.

The most important property – necessary for assuming objectively good cause-and-effect hypotheses – seems to be that the person has a profound knowledge about actions carried out in the past and their implications for the company's measures. As a consequence the proposed approach provides a tool to make these facts explicit in the form of a model and verify the beliefs, underlying a causal influence, by the means of an Artificial Neural Network.

For this purpose one has to compare the similarity between different patterns that (may) occur within the relevant samples of incoming and resulting indicators of an influence-relation. As mentioned in the previous section, there are several statistical techniques (see also [9] and [12]) and applications of Artificial Intelligence to solve pattern recognition problems. Since this specific issue is characterized by a considerable amount of so-called white noise resulting from the (unknown) influence of exogenous indicators (i.e. variables not included in the model), a procedure has to be chosen which can tackle this problem: Artificial Neural Networks have the favorable feature to deal with noisy input by massively parallel operations carried out by neurons and consequently seem to be a proven tool for the verification of cause-and-effect assumptions. Furthermore this technique can be regarded as universal function approximators in the sense of HECHT-NIELSEN´s function approximation networks as it has been shown in the previous section.

In order to apply Neural Network techniques to the already proposed modeling environment the meta-model is extended as follows:

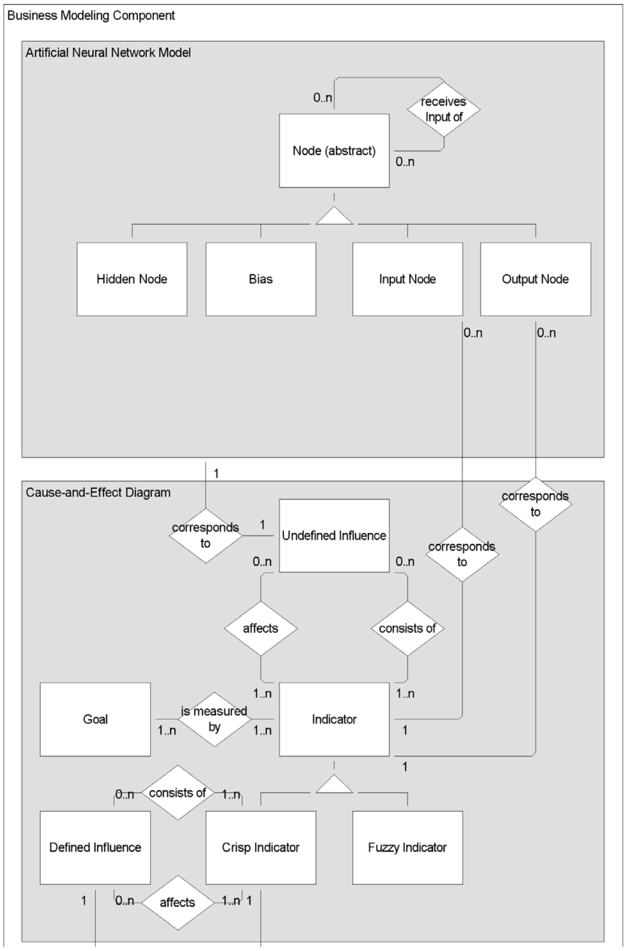

figure 3: Detail of the extended meta-model

As depicted in figure 3 the relation "Undefined Influence" is raised to the state of a class, which is linked to incoming indicators using the relation "consists of". The outgoing relation of the type "affects" links the "Undefined Influence"-class with the result measure. After decision-makers have identified sufficient unknown influences affecting a specific indicator, they estimate the magnitude of this relation by rating its share of the total influence on this indicator as a percentage.

An affiliation graph, which is constructed by the relations connected to an instance of the class "Undefined Influence", is an important prerequisite for building an Artificial Neural forecasting-Network: On request, a customizable procedure within the decision supporting system constructs automatically a feedforward neuron model containing input nodes which correspond to the incoming indicators of the unknown influence and an output node equivalent to the result measure. As a consequence a standard procedure is designed to construct a three-layer feedforward perceptron. Since there exist numerous heuristics for the specification of an appropriate network architecture, these can be used to avoid overtraining of the resulting MLP: The following heuristic is used to identify an optimal dimension of the hidden layer (determined by the number of training patterns N_{TP} and the size of the input layer $dim(I^{in})$):

$$\dim(Hidden) = \left\lceil \frac{N_{TP}}{5 \cdot \dim(I^{in}) + 1} \right\rceil \tag{6}$$

An instance of the model type "Artificial Neural Network Model" is created, where the neuron model – constructed as described above – is stored.

At this stage, training data sets can be derived out of the operational databases by means of a data extraction system (described in [15]): Historic time series of the input indicators provide input patterns for the reproduction of the perceptron's output values. These are compared with the corresponding expected results (derived from the output indicators) and end up in the calculation of a reproduction error.

After the termination of this translation procedure, the training of the connectionist system is carried out. Therefore the back-propagation algorithm is applied to the ANN-model in the development environment for Artificial Neural Networks - VIENNA (see [11] for a detailed discussion). Following ROSENBLATT's perceptron convergence theorem [24], self-adapting learning stepwidths and momentum factors ensure that the training algorithm is approximating a global error minimum. This least remaining deviation can be used to make a statement about the degree of convergence of the neural model and – as a consequence – it can be taken as an indicator for the correlation between the incoming and resulting time series.

Setting the remaining errors of the perceptrons (corresponding to the unknown influences) into relation, one can examine the fitness of the assumptions made before, by comparing the decision-maker's estimates with the Objective Shares of Influence, which are calculated as shown below:

$$OSI_i = \frac{\dfrac{RE_i}{\max(RE_n)}}{\sum_{j=1}^{J} \dfrac{RE_j}{\max(RE_n)}} \tag{7}$$

where OSI_i is the Objectve Share of Influence and RE_i the Remaining Error of indicator i.

6. SUMMARY, CONCLUSIONS AND FUTURE WORK

This paper proposes an approach to model and evaluate the complex causal relations, which have to be analyzed for strategic planning issues. The rudimentary cause-and-effect diagram has been extended in order to represent known influences between crisp parameters, which are characterized by axiomatic construction rules. Some managerial theories as the Balanced Scorecard approach [18] and the Sensitivity Model [34] assume that there is a negative association between the activity of an indicator (i.e. its impact on the overall system) and its ability to be measured by numeric means. Since many Decision Support Systems based on causal concepts only assume hypothetical belief structures as a foundation for building strategies, this approach pursues the intention to infer unknown causal relations from empirical observations: Transforming a single-stage influence relation into a multi layer perceptron with training patterns derived from operational databases, it is possible to automatically ascertain the causality function of influencing and influenced indicators by the means of the smallest Remaining Error. The ability of this approach to reconstruct causal structures within business models as well as to approximate the underlying cause-and-effect function at a sufficient accuracy level has been proven in an artificial environment consisting of synthetically generated causal timeseries (cf. [14]). Therefore a future research issue will be to research these abilities within a real-life industrial setting.

7. REFERENCES

[1] S. Beer, "Brain of the firm – The managerial cybernetics of organization", John Wiley & Sons, 2nd edition, 1981.

[2] H. Brightman and H. Schneider, "Statistics for business problem solving", South-Western Publishing, 2nd edition, 1994.

[3] P. R. Cohen, L. A. Ballestros, D. E. Gregory, R. S. Amant, "Regression can build predictive causal models", Computer Science Technical Report 94-15, Experimental Knowledge Systems Laboratory, University of Massachusets, Amherst (MA), 1994

[4] G. Cooper and E. Herskovits, "A bayesian method for constructing bayesian belief networks from databases", in B. D'Ambrosio, P. Smets and P. Bonissone (eds.), "Uncertainty in Artificial Intelligence", pp. 86-94, Morgan Kaufmann, San Francisco (CA), 1991.

[5] J. Denton, "How good are neural networks for causal forecasting?", Journal of Business Forecasting (Summer 1995), pp. 17-20.

[6] B. Gerstman and M. Innovera, "StatPrimer, Version 5.1. An on-line, public-domain instructional manual", http://www.sjsu.edu/faculty/gerstman/StatPrimer/ (downloaded: 03-13-2002).

[7] S. Gonzalez, "Neural networks for macroeconomic forecasting: A complementary approach to linear regression models", Working Paper 2000-07, Canadian Department of Finance, 2000

[8] R. Hecht-Nielsen, "Kolmogorov's mapping neural networks existence theorem", in: "Proceedings of the IEEE international conference on neural networks", vol. III, pp. 11-14, IEEE-Press, New York (NY), 1987

[9] C. Hillbrand, "Corporate Performance Management: Strategische Planung durch ganzheitliche Unternehmenskennzahlen", in B. Britzelmaier and S. Geberl (eds.), "Information als Erfolgsfaktor", Teubner Verlag Reihe Wirtschaftsinformatik, Stuttgart – Leipzig – Wiesbaden, 2000, pp. 211-220 (in German).

[10] C. Hillbrand, "Management support systems", Technical Report No. 01/2001, University of Vienna, Vienna, Austria, May 2001 (in German).

[11] C. Hillbrand, "Einsatz konnektionistischer Systeme in betriebswirtschaftlichen Modellen zur Unterstützung strategischer Entscheidungen", Technical Report No. 02/2001, University of Vienna, Vienna, Austria, September 2001 (in German).

[12] C. Hillbrand, "Methoden zur Bewertung von Kausalzusammenhängen in betriebswirtschaftlichen Modellen", Technical Report No. 01/2002, University of Vienna Wien, Vienna, Austria, April 2002 (in German).

[13] C. Hillbrand, "Inferenz in betriebswirtschaftlichen Kausalmodellen auf der Basis von Ursache-Wirkungshypothesen", Technical Report No. 02/2002, University of Vienna, Vienna, Austria, November 2002 (in German).

[14] C. Hillbrand, "Verifikation von betriebswirtschaftlichen Ursache-Wirkungshypothesen durch kausale Inferenz", Research Paper, University of Vienna, Vienna, Austria, February 2003 (in German).

[15] C. Hillbrand and D. Karagiannis, "Using artificial neural networks to prove hypothetic cause-and-effect relations: A meta-model-based approach to support strategic decisions", in M. Piattini, J. Filipe and J. Braz, "Proceedings of ICEIS 2002 - The Fourth International Conference on Enterprise Information Systems", vol. 1, pp. 367-373, Universidad de Castilla-La Mancha, Ciudad Real, Spain, April 3-6, 2002.

[16] P. Holland, "Statistics and causal inference", Journal of the American Statistical Association, vol. 81, pp. 945-960, 1986.

[17] Y. Iwasaki and H. A. Simon, "Causality and device behavior", Artificial Intelligence, vol. 29, pp. 3-32, 1986.

[18] R. Kaplan and D. Norton., "The balanced scorecard: Measures that drive performance", Harvard Business Review, vol. 70, no. 1, January-February 1992, pp. 71-79.

[19] R. Kaplan and D. Norton, "Using the balanced scorecard as a strategic management system", Harvard Business Review, vol. 74, No. 1, January-February 1996, pp. 75-85.

[20] D. Karagiannis and R. Telesko, "Wissensmanagement: Konzepte der Künstlichen Intelligenz und des Softcomputing", Oldenbourg Verlag, München – Wien, 2001 (in German).

[21] J. Pearl, "Causality: Models, reasoning and inference", Cambridge University Press, Cambridge – New York - Melbourne – Madrid, 2000.

[22] J. Pearl and A. Paz, "Graphoids: A graph-based logic for reasoning about relevance relations", in B. du Bolay et al. (eds.), "Advances in artificial intelligence", vol. II, pp. 357-363, North Holland, Amsterdam, 1987.

[23] J. Pearl and T. S. Verma, "A theory of inferred causation", in J. Allen , R. Fikes and E. Sandewall (eds.), "Principles of knowledge representation and reasoning: Proceedings of the second international conference", pp. 441-452, Morgan Kaufmann, San Francisco (CA), 1991.

[24] F. Rosenblatt, "Principles of neurodynamics", Spartan, New York (NY), 1962.

[25] W. S. Sarle, "Neural networks and statistical models", in "Proceedings of the Nineteenth Annual SAS Users Group International Conference", April 1994.

[26] H. Simon, "Spurious correlations: A causal interpretation", Journal of the American Statistical Association, vol. 49, 1954, pp. 469-492.

[27] A. E. Smith and A. K. Mason, "Cost estimation predictive modeling: Regression versus neural network", The Engineering Economist, vol. 42, no. 2, Winter 1997, pp. 137-161.

[28] P. Spirtes, C. Glymour and R. Scheines, "Causation, prediction and search", Springer Verlag, New York (NY), 1993.

[29] S. S. Stevens, "On the theory of scales of measurement", Science 161, 1946, pp. 677 – 680.

[30] F. M. Thiesing and O. Vornberger, "Sales forecasting using neural networks", in "Proceedings of ICNN", vol. 4, pp. 2125-2128, IEEE, Houston (TX), 9-12 June 1997.

[31] D. Tikk, L.T. Kóczy and T.D. Gedeon, "A survey on the universal approximation and its limits in soft computing techniques", International Journal of Approximate Reasoning, vol. 33, no. 2, 2003, pp. 185-202.

[32] F. Vester, "The biocybernetic approach as a basis for planning our environment", Systems Practice, vol. 4, pp. 399-413.

[33] F. Vester, "Simulating complex systems as sustainable organization by transparent sensitivity models", Lecture at Eurosim 2001 Congress, 21-29 June, Delft, 2001.

[34] F. Vester, "Sensitivity model Prof. Vester", http://www.frederic-vester.de/model.htm. (downloaded 03/20/2002)