European Journal of Operational Research 130 (2001) 623–637

EUROPEAN JOURNAL OF OPERATIONAL RESEARCH

www.elsevier.com/locate/dsw

Theory and Methodology

A general model for automated business diagnosis

A.J. Feelders [a,*], H.A.M. Daniels [a,b]

[a] Department of Economics and Business Administration, Center for Economic Research, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, Netherlands

[b] Rotterdam School of Management, Institute of Advanced Management Studies, P.O. Box 1738, 3000 DR Rotterdam, Netherlands

Received 3 June 1998; accepted 7 September 1999

Abstract

In this paper we develop a formalism to support diagnostic reasoning in the domain of business and finance. A theoretic description of the process of diagnosis of company performance, and the implementation thereof is outlined. A new concept of explanation comprises the basis of our framework and enables us to deal with both qualitative and quantitative information in diagnosis. The system was tested on a case-study involving the comparison of nine firms operating in the mechanical engineering industry. Comparison of textbook analysis and model output shows that our system is able to produce the correct diagnostic analyses. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Management; Decision support systems; Finance; Artificial intelligence

1. Introduction

The formalisation of diagnostic problem solving is a sub-area of Operations Research and Artificial Intelligence (AI) that has received considerable attention in recent years. Diagnosis is usually defined as finding the best explanation of observed abnormal behaviour of a system under study. Especially *model based*, as opposed to *heuristic classification*, approaches have been developed and investigated, mainly because of their supposed superior problem solving and explanation capabilities.

The larger part of research into diagnostic problem solving has either implicitly or explicitly been concerned with medical diagnosis or diagnosis of man-made artifacts such as electronic circuits. This has had consequences for the applied knowledge representation formalisms and associated reasoning methods. For example, in the domain of diagnosis of electronic circuits, the system concerned is usually represented as a set of first-order logic sentences, describing the circuit components and the way they are interconnected. In the medical domain one often encounters causal models, that describe cause–effect relations between *disease states*. In accordance with the knowledge available in the domain these causal models are of a qualitative nature. We take a causal view of explanation that is able to deal

* Corresponding author. Tel.: +31-13-466-8201; fax: +31-13-466-3377.

E-mail address: A.J.Feelders@kub.nl (A.J. Feelders).

PII: S 0 3 7 7 - 2 2 1 7 (9 9) 0 0 4 2 8 - 2

with both the qualitative and quantitative phenomena that pervade the domain of business and finance.

With respect to the issue of knowledge representation, our formalisation can be characterised as a *model based* rather than a *heuristic rules* approach. Apart from the advantages ascribed to a model based approach, it also appears to be the most natural approach in the domain of business and finance. On the one hand this is the case because a substantial part of the knowledge involved deals with relations between financial and operational variables. This knowledge is usually already available in the form of a system of equations that are used to *compute* values rather than to *explain* them. On the other hand there is knowledge of cause–effect relations that is often of a qualitative nature [5].

This paper is organised as follows. In Part I we develop the theoretical underpinnings of our diagnostic method. In Section 3, we describe a causal model of explanation, which serves as the basis for diagnostic reasoning. In Section 4 we discuss the representation of knowledge in the business model. In Section 5 we present the diagnostic method. We conclude Part I with an overview of related work. In Part II (which consists solely of Section 8) the method is illustrated in a case study on interfirm comparison. Finally, we draw a number of conclusions in Section 9.

Part I: Theory

2. Motivation

Before we discuss the theoretical underpinnings of our method, we first provide a little example to show what motivated us to develop it in the first place. This example is discussed more elaborately in Part II, Section 8.

In [19] (p. 35) the following explanation is given for a firm's (indicated in the text as firm 1) relatively, i.e. compared to other firms, high return on operating assets (ROA):

> *Why is firm 1's return on operating assets relatively high? Comparisons of ratios 2 (profit margin) and 3 (turnover of assets) with those of the other firms show that the firm's high ratio 1 (return on operating assets) is due to a combination of a comparatively high profit on sales and a comparatively fast turnover of assets.*

The analysis then continues with the explanation of the contributing factors identified. This kind of reasoning seems essential in analysing and explaining company performance. The analyst notices that firm 1 has a value for "return on operating assets" that deviates substantially from some "norm value", and attempts to provide an explanation in terms of other variables that contributed to this deviation.

Our aim is to identify the knowledge structures and reasoning methods required to construct such explanations from data on company performance. These formal methods may then be used to develop a computer algorithm to support human decision makers in the diagnostic process.

3. A causal model of explanation

According to a causal model of explanation, we explain phenomena by giving their causes. Our exposition on causal explanation is largely based on Humphreys' notion of *aleatory* explanations ([17]). Aleatory explanations have been introduced to account for probabilistic phenomena, but they are also applicable in deterministic contexts. Causal influences come in two kinds: *contributing* and *counteracting*. Therefore, Humphreys proposes the following canonical form for causal explanations:

Event E occurred because of C^+, despite C^-,

where E is the event to be explained, C^+ is a non-empty set of contributing causes, and C^- a (possibly empty) set of counteracting causes. The explanation itself consists in the causes to which C^+ jointly refers. C^- is not a part of the explanation of E, but gives us a clearer notion of how the members of C^+ actually brought about E. Thus C^- may be empty, in which case we have an explanation involving only contributing causes to E's occurrence, but if C^+ is empty, then we have no explanation of E's occurrence at all.

We did not yet thoroughly specify what is to count as an event in the context of causal explanations. In company diagnosis, one can at first sight identify two kinds of events:

- variable y has some particular value at time t,
- variable y *changes* value from time t to t'.

An explanation of the first kind has an empty "despite" clause, since all causally relevant factors will *contribute* to y having that exact value. We claim that this kind of explanation is not very interesting for diagnostic purposes, and therefore we will ignore it in this paper.

We will generalize the second type of event, in order to be able to explain a broader class of phenomena. The event of a variable y changing from time t to t' can be shown to be a special case of the more general "event" of there being a difference between two values of y evaluated with respect to different "objects". Since we are interested in explaining the difference between actual and normal behaviour in diagnosis, this type of explanation meets our goals. In order to make this clear we discuss a theory of explaining differences as developed by Hesslow [16].

According to Hesslow all explanations of individual facts of the form $F(a)$ – object a has property F – involve a, sometimes implicit, comparison with other objects which lack the property in question. The objects of comparison are said to belong to a reference class R. The only restriction put on R is that its members must not possess the explanandum property. Hesslow's theory leads to a more detailed specification of the event to be explained. Instead of representing the explanandum as $F(a)$, it is specified further by explicitly including the reference class R. Consequently, the explanandum is a three-place relation $\langle a, F, R\rangle$ between an object a (e.g. the ABC-company), a property F (e.g. having a relatively low return on assets) and a reference class R (e.g. other companies in the industry). The task is not to explain why a has property F, but rather to explain why a has property F *when the members of R do not.* There are several typical ways of forming reference classes, for example:

(a) *R as the statistically normal case*. If a certain causal condition were normal, it would occur among the members of R. In that case it could never explain the difference between the explanandum object and those in R. It follows that when R is chosen as the statistically normal, the explanatory cause must be abnormal.

(b) *R as the temporally normal case*. If the question is "why did profit of the ABC-company decrease at this particular time", we are asking about a temporal difference, and the proper object of comparison will be the ABC-company at an earlier time.

(c) *R as a theoretical ideal*. In many sciences it is a common procedure to use as an object of comparison a *hypothetical* object or state of affairs which is defined by some theory. Such *theoretical ideals* have the obvious advantage of providing the scientist with a constant object of comparison, thus facilitating systematization of the field covered by the theory. A typical example of such a theoretical ideal in medicine is the physiology of the healthy human organism.

For the purpose of explanation, the class R can often be reduced to *one* member r, which is in some sense the average of the class R or the ideal object.

If E is replaced by this more detailed explanandum, the following new canonical form for explanations is obtained:

$\langle a, F, r\rangle$ because C^+, despite C^-.

In this section we described the structure of an explanation. It is generally agreed upon that explanations of events should be based on "general laws". In the next section we discuss the types of general laws that we think are relevant for the domain of business and finance.

4. The business model

Explanations are usually based on general laws expressing relations between events, such as cause–effect relations or constraints between variables. The general laws on which explanations are based, are represented in the business model M. The business model M represents relevant financial and operating variables and the relations between them. We distinguish between quantitative and qualitative relations among model variables:

(a) $y = f(x_1, \ldots, x_n)$,
(b) $y \leftarrow \{(x_1, sign_1), \ldots, (x_m, sign_m)\}$,
where
$sign_i \in \{\text{pos}, \text{neg}\}$.

Quantitative relations are for example used to represent definitions, such as profit = revenues − costs, or consolidation equations. Quantitative relations can also be used to represent causal relations, determined by econometric analysis. In most cases however, knowledge of causal relations will only be available in qualitative form. The second type of relation expresses knowledge where the exact quantitative relation between variables is not known. For example, "an increase in competition will lower the selling price" is represented as follows:

$$\text{selling price} \leftarrow \{(\text{competition}, \text{neg})\}.$$

The relevance of qualitative knowledge for economic analysis has been confirmed by a number of studies [4,18,2]. Such qualitative relations are bipolar, i.e. an increase in competition leads *ceteris paribus* to a price decrease, and a decrease in competition to an increase in price. The qualitative part of the business model is similar to well-known qualitative modelling approaches such as used in cognitive mapping [11] and signed digraphs [15].

It is important to note here that the business model has a "mixed ontology". Quantitative functional relations refer to a state of the system at a particular point in time, whereas qualitative relations refer to *differences* between states. Quantitative model relations are not required to be additive, because non-additive equations are quite common in financial models. We do put some restrictions on the quantitative equations however; for a discussion of these restrictions we refer to Section 6.

Both types of relations are interpreted in such a way that the one variable that appears on the left-hand side of a relations is *caused by* or *depends on*

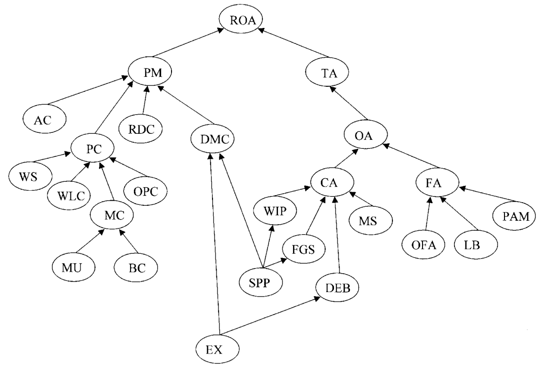

Fig. 1. Explanatory graph for interfirm comparison model.

the variables on the right-hand side. With the business model M we associate a directed graph $E(M) = (\mathscr{V}, \mathscr{E})$, called the *explanatory* graph of M, as follows. The vertex set $\mathscr{V}$ contains as elements all variables appearing in the model. The edge set $\mathscr{E}$ contains a directed edge from vertex x_i to x_j iff:

1. $x_j = f(\ldots, x_i, \ldots) \in M$, or
2. $x_j \leftarrow \{\ldots, (x_i, sign_i), \ldots\} \in M$.

A restriction we put on the model M is that its explanatory graph $E(M)$ should not contain any cycles, since this would make explanations circular. This restriction excludes models that contain simultaneous equations. Nodes in the explanatory graph, with zero indegree, represent variables that cannot be explained in M. The explanatory graph of the business model in Section 5, is depicted in Fig. 1.

In the next section we provide an operational definition of the concept of explanation, using both the quantitative and qualitative relations of the business model. This will lead to the development of novel reasoning methods for the diagnosis of company performance.

5. Diagnosis and explanation

In this section we propose a method of explanation for the diagnosis of company performance, that is based on the canonical explanation format presented in Section 3. For the purpose of diagnosis of company performance, we are interested in explaining the difference between the actual and norm behaviour of a particular business company. The event $\langle a, F, r\rangle$ to be explained is now specified as follows:

1. $a =$ the actual behaviour of a company,
2. $F =$ a particular variable deviates from its norm value,
3. $r =$ the norm behaviour for the company involved.

Since the object a and reference object r will always be clear from the context, we simplify the *canonical explanation* format to

$$\partial y = q \text{ occurred because } C^+, \text{ despite } C^-.$$

In this expression, $\partial y = q$ specifies the event, i.e. the occurrence of qualitative difference between the *actual* and *norm* value of y, denoted by y^a and y^r, respectively.

This qualitative difference can take on one of the values {low, normal, high}, and is determined according to the rules of Table 1. For the purpose of diagnosis, we are not interested in explaining $\partial y =$ normal, since it is only required to explain why a variable deviates from its norm value. Note furthermore, that an explanation for $\partial y =$ normal, would actually be an explanation for a quantitative difference, i.e. it would explain why $\Delta y = y^a - y^r = 0$.

The way in which contributing and counteracting causes are determined depends on the type of relation, quantitative or qualitative, from the business model that sustains the explanation. First we discuss the situation where the explanation is sustained by a quantitative equation $y = f(\mathbf{x})$ where $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ denotes an n-component vector. To determine the contributing causes (C^+) and counteracting causes (C^-) that explain the qualitative difference between the actual and norm value of y, we define a measure of influence as follows (cf. [22,13]):

Table 1
Mapping to qualitative value

	∂y
$y^a > y^r$	High
$y^a = y^r$	Normal
$y^a < y^r$	Low

$$\inf(x_i, y) = f(\mathbf{x}^r_{-i}, x^a_i) - y^r,$$

where $f(\mathbf{x}^r_{-i}, x^a_i)$ denotes the value of $f(\mathbf{x})$ with all variables evaluated at their norm values, except x_i. In words, $\inf(x_i, y)$ indicates what the difference between actual and norm value of y would have been if *only* x_i would have deviated from its norm value.

The influence of a set of variables is a straightforward extension of the influence of a single variable. The influence of a set of variables X with index set $I_X \subseteq \{1, 2, \ldots, n\}$ on y is defined as

$$\inf(X, y) = f(\mathbf{x}^r_{-I_X}, \mathbf{x}^a_{I_X}) - y^r,$$

where $f(\mathbf{x}^r_{-I_X}, \mathbf{x}^a_{I_X})$ denotes the value of $f(\mathbf{x})$ that results from evaluating all components with index in I_X at their actual values, and all other components of $\mathbf{x}$ at their norm values.

Determination of the influence of x_i on y involves the evaluation of a hypothetical situation represented by the expression $f(\mathbf{x}^r_{-i}, x^a_i)$. The correct interpretation of the inf-measure depends, as was to be expected, on the form of the function f. Let us assume for the moment that previous period values serve as norm values, so we are effectively explaining a change of y from period $t-1$ to t. If f is additive, then $\inf(x_i, y)$ is correctly interpreted as a quantitative specification of the change in y that is explained by the change in x_i from $x_{i,t-1}$ to $x_{i,t}$. If f is monotonic but non-additive, then the interpretation of $\inf(x_i, y)$ is considerably more difficult. It can certainly *not* be interpreted as a quantitative specification of the change in y that is explained by the change in x_i alone, since the value of $\inf(x_i, y)$ also depends on the $t-1$-level at which variables other than x_i are held constant. Hence, such a quantitative claim only holds true within that particular context. However the sign of $\inf(x_i, y)$ is not context-dependent for monotonic functions (see Section 6 for further elaboration on these issues).

We define contributing and counteracting causes as follows.

Definition 1 (*Contributing* (*Counteracting*) *causes*). The set of contributing (counteracting) causes

C^+ (C^-) consists of the components x_i of $\mathbf{x}$ with $\inf(x_i, y) \times \Delta y > 0$ (< 0).

As currently defined, the explanation method does not possess the desired property of leaving insignificant influences out of the explanation. This could potentially lead to an *information overload*. It is therefore not surprising that observation of human financial analysts shows that they tend to leave out insignificant influences from the explanation. To obtain the desired behaviour we define parsimonious sets of contributing (counteracting) causes as follows:

Definition 2 (*Parsimonious set of contributing causes*). The parsimonious set of contributing causes C_p^+, is the smallest subset of C^+ such that

$$\frac{\inf(C_p^+, y)}{\inf(C^+, y)} \geqslant T^+.$$

The parsimonious set of counteracting causes is defined analogously. The parsimonious set of contributing causes is the smallest subset of the set of contributing causes, such that its influence on y exceeds a particular fraction (T^+) of the influence of the complete set. In case there are several sets of equal cardinality that explain a fraction larger than T^+, the one with the highest inf-value is called the parsimonious set. The definition with respect to counteracting causes is clearly analogous.

The fractions T^+ and T^- are numbers between 0 and 1, and will typically be close to 1. In the sequel we use the following format for parsimonious one-level explanations:

$\partial y = q$ because C_p^+, despite C_p^-.

Thus far, we have been concerned with explanations based on quantitative equations. Now we direct our attention to explanations based on qualitative relations. A qualitative relation only states whether a variable has a positive (pos) or negative (neg) influence on another variable. Since the strength of such an influence is not quantified, we only consider the *qualitative* difference between actual and norm value when determining contributing and counteracting causes. For example, if x has a positive influence on y and $\partial x =$ low, and $\partial y =$ low, then x is a contributing cause of $\partial Y =$ low. On the other hand, if x were to have a negative influence, then it would be a counteracting cause.

In case of explanations based on qualitative relations the distinction between complete and parsimonious explanations disappears, since there is no possibility to weigh the different contributing and counteracting causes. Therefore, $C_p^+ = C^+$ and $C_p^- = C^-$.

Thus far, we have discussed *one-level* explanations. These explanations are one level deep, in the sense that they are based on only one relation from the business model. It is useful for diagnostic purposes however, to continue an explanation of $\partial y = q$, by explaining the qualitative difference between the actual and norm values of its contributing causes. This process can be continued until a contributing cause is encountered that cannot be explained within the business model, because the business model does not contain a relation in which this contributing cause appears on the left-hand side. More generally, we define a maximal explanation for $\partial y = q$ as follows.

Definition 3 (*Maximal explanation*). A maximal explanation for $\partial y = q$ is a tree with the following properties:
1. y is the root node of the tree.
2. The root node has two types of successors, corresponding to its parsimonious contributing and counteracting causes, respectively.
3. A node that corresponds to a contributing cause has two types of successors, corresponding to its parsimonious contributing and counteracting causes, respectively.
4. A node that corresponds to a counteracting cause has no successors.
5. A node that corresponds to a variable that cannot be explained in the business model has no successors.

In this definition, the different status of contributing and counteracting causes is clearly reflected. Without contributing causes there would be no explanation at all, whereas the absence of counteracting causes merely indicates that there was no opposition against $\partial y = q$. Therefore, a maximal explanation for $\partial y = q$ continues with its contributing causes, whereas the counteracting causes are not explained any further.

A diagnosis is an explanation for observed abnormal behaviour, sometimes called problem identification, of a company. For problem identification we use a norm model that contains a norm unit for a subset of the variables in the business model, called the *performance indicators*. These indicators typically correspond to critical financial and operational measures such as used in methods like the balanced scorecard [21].

Problem identification is a fairly straightforward process that computes a value $g(y^a, y^r)$ for each performance indicator, where g is some user-specified function such as percentage difference or absolute difference. If this value is below some specified threshold, a symptom $\partial y = \text{low}$ is added to the list of symptoms. Likewise, if the value of g is above some specified upper threshold, a symptom $\partial y = \text{high}$ is added. The result of problem identification is a set of symptoms $S = \{\partial y_1 = q_1, \ldots, \partial y_n = q_n\}$, where $q_i \in \{\text{low}, \text{high}\}$. We define the data set Z, as the set that contains all actual values and norm values, for the variables in M. A complete diagnosis for the triple $\langle M, Z, S\rangle$ is then simply the set of maximal explanations in business model M for all symptoms in S, using data set Z.

6. Restrictions on quantitative equations

The diagnostic method presented in Section 5 assumes that it makes sense to determine for each variable x_i separately whether it had a positive or negative influence on y. Now consider the example in Table 2.

Table 2
Actual and norm values for $y = x_1 \times x_2$

	r	a	$\inf(x_i, y)$
x_1	5	−8	−78
x_2	6	−5	−55
y	30	40	10

From the last column of this table we infer that x_1 and x_2 when considered separately both have a negative influence on y, but their joint influence $y^a - y^r$ is positive. The influence of x_1 (x_2) changes sign when considered in conjunction with x_2 (x_1). Only x_1 and x_2 together can explain the increase in y and an attempt at decomposition of their contributions leads to counterintuitive results.

The so-called conjunctiveness constraint captures the intuitive notion that the influence of a single variable should not turn around when it is considered in conjunction with the influence of other variables. Only when this constraint is satisfied, can we say for each variable separately whether it had a positive or negative influence, and therefore whether it was a contributing or counteracting cause. Furthermore, satisfaction of this constraint guarantees the existence of an efficient algorithm for explanation generation [12].

An equation satisfies the conjunctiveness constraint if for all subsets $X \subseteq \{x_1, \ldots, x_n\} \setminus \{x_i\}$:
1. $\inf(x_i, y) > 0 \Rightarrow \inf(X \cup \{x_i\}, y) > \inf(X, y)$,
2. $\inf(x_i, y) < 0 \Rightarrow \inf(X \cup \{x_i\}, y) < \inf(X, y)$,
3. $\inf(x_i, y) = 0 \Rightarrow \inf(X \cup \{x_i\}, y) = \inf(X, y)$.

In this section we apply some simple algebraic transformations to the model to determine what kind of equations satisfy the conjunctiveness constraint. First define $g : [0,1]^n \to \mathbb{R}$ by

$$g(\mathbf{s}) = \frac{f(\mathbf{s} \cdot \mathbf{x}^a + (\mathbf{1} - \mathbf{s}) \cdot \mathbf{x}^r) - y^r}{y^a - y^r}. \tag{1}$$

Here $\mathbf{s} = (s_1, \ldots, s_n) \in [0,1]^n$, $\mathbf{x}^a = (x_1^a, \ldots, x_n^a)$ and $\mathbf{x}^r = (x_1^r, \ldots, x_n^r)$.

The function g is defined on the unit hypercube and $g(0, \ldots, 0) = 0$, $g(1, \ldots, 1) = 1$. The conjunctiveness constraint reformulated in terms of the function g states that for each direction i with $g(\mathbf{1}_i) > 0$ and vertex $\mathbf{v}$ of the unit hypercube with $v_i = 0$:

$$g(\mathbf{1}_i) \cdot \{g(\mathbf{v} + \mathbf{1}_i) - g(\mathbf{v})\} \geqslant 0,$$

where $\mathbf{1}_i$ denotes the vector $(v_1, \ldots, v_n)$ with $v_i = 1$ and $v_j = 0$ for $j \neq i$.

The following lemmas characterize a large class of models that satisfy the conjunctiveness constraint.

Lemma 1. *If f is additive, then f satisfies the conjunctiveness constraint.*

Proof. Note that additivity is preserved by the transformation (1), and therefore

$$g(\mathbf{1}_i) \cdot \{g(\mathbf{v}+\mathbf{1}_i) - g(\mathbf{v})\} = g(\mathbf{1}_i) \cdot g(\mathbf{1}_i) \geqslant 0$$

since $g(\mathbf{1}_i) > 0$. □

Also monotonic functions satisfy the conjunctiveness constraint. By monotonicity we mean monotonicity in all variables separately, on the domain under consideration.

Lemma 2. *If f is monotonic then f satisfies the conjunctiveness constraint.*

Proof. The proof is given for the case that f is differentiable (if f is not differentiable, the algebra is a little more complicated). Note that monotonicity of f implies that the sign of all partial derivatives of g is constant on the unit hypercube, i.e. $\text{sign}(\partial g/\partial s_i)$ is constant. We have $\text{sign}(g(\mathbf{1}_i)) = \text{sign}(\partial g/\partial s_i)$ and also $\text{sign}(g(\mathbf{v}+\mathbf{1}_i) - g(\mathbf{v})) = \text{sign}(\partial g/\partial s_i)$. Hence

$$g(\mathbf{1}_i) \cdot \{g(\mathbf{v}+\mathbf{1}_i) - g(\mathbf{v})\} \geqslant 0. \qquad \square$$

It turns out that in practice one often encounters models that are combinations of additive and monotonic functions. Additivity and monotonicity can also be easily verified. For example in the model described in Section 8 the quantitative part consists of 2 monotonic relations and 5 additive relations.

7. Related work

In this paper we focus on knowledge representation for business diagnosis and methods for diagnostic inference. The method proposed here is based on a hybrid framework combining quantitative and qualitative models and reasoning. We discuss related work for the quantitative and qualitative part separately, since most systems address either one or the other, but not both. We also briefly consider some approaches that may prove helpful aids in diagnostic problem solving, but are not directed at full automation of this process.

7.1. Qualitative component

The qualitative part of the method is related to earlier work published by researchers working in a wide variety of application areas. For example in [10], cognitive maps serve as representations of managers thinking about business processes. Qualitative relations are represented as signed digraphs. The authors of [10] have also implemented a program called COPE for automatic path analysis. This work is similar to the method described in Section 5. For example, the COPE system allows the user to explore *possible* explanations of an increase or decrease in a particular variable of interest. Our diagnostic algorithm also determines however which of the possible explanations *actually* apply to the specific situation considered, based on the available data.

In our framework we only consider monotonic relations as they are called in [10]; we also restrict to models without feedback loops. The method of cognitive maps of Eden et al. is mainly used to describe a problem situation as perceived by an individual manager or a group of managers. It supports learning about the problem and problem environment and is also used for policy exploration and evaluation. Our objective is to use the qualitative model in combination with the quantitative part for formal diagnostic reasoning by means of a computer program.

In this respect our approach is also related to the work of Bouwman [4,3] who studied the diagnostic reasoning of experienced financial analysts and compared this to the problem solving behaviour of novices. He also developed computer programs that can mimic the behaviour of human analysts including the shortcomings and mistakes that occurred in their analyses. The approach of Bouwman is descriptive in contrast to the normative line followed in this paper.

There is also a vast amount of work in qualitative models for physical systems as well as diagnosis of physical systems based on qualitative models [25]. The approach followed in qualitative physics is rather different from the methodology in business administration and economics, since both the relationships (confluences) and the way of reasoning applied to them (causal heuristics) has no counterpart in economics. In economics qualitative models have been applied to simulate the future behaviour of dynamic economic systems, such as markets or economies consisting of coupled markets [2].

7.2. Quantitative component

Several approaches have been proposed for the automatic generation of explanations based on quantitative (financial) models in the literature [1,8,22,23].

Kosy and Wise [22,23] describe a general system for generating explanations in financial models, not directed specifically at diagnostic problem solving. Their algorithm can explain any difference between two values of a variable, as long as these values have been generated by the same equation. They do not make a strict separation between contributing and counteracting causes however, which leads to counterintuitive results in some cases. Most important is that it may cause the system to leave out significant contributing or counteracting causes from the explanation.

Courtney et al. [1,8] describe a decision support system directed specifically at managerial problem diagnosis. Functional relations that are allowed to sustain explanations are restricted to *linear* functions however. This avoids the difficulties with respect to making quantitative causal claims that occur when explanations are based on non-additive functions. The restriction to linear relationships is not very realistic however in a financial context. A clear distinction is made in their system between contributing and counteracting influences. The system is not fully automatic however, in the sense that the user has to decide which of the complete list of influences presented is considered relevant.

There has been quite some attention for research on the assessment of business performance, such as Data Envelopment Analysis (DEA) [6,7] and the balanced scorecard [21]. DEA is a powerful and widely applicable technique for assessing the relative efficiency of a number of decision making units based on the comparison of achieved output(s) at particular levels of inputs. The efficiency of the units is determined by solving a linear programming problem. One of its strong points is that all relevant outputs can be considered simultaneously.

Our aim here is to look inside the black box to show the relation between particular symptoms and other relevant variables. The intermediate (not input or output) variables occur in an explanation tree generated by the diagnostic program. The main difference between DEA and the method proposed here is that we look into further detail to establish causes for symptoms, whereas DEA investigates overall efficiency. To apply our diagnostic procedure, not only input/output relations are required, but a model describing the multi-level dependency structure between relevant operational and financial variables.

The balanced scorecard is a method for performance assessment of companies, based on a collection of both financial and operational performance measures. As is recognized in such methods, bad performance in a particular area may be part and parcel of the company's policy. The method proposed in this paper allows us to uncover such relationships by systematic generation of explanations for the observed symptoms.

Another related branch of research is concerned with the *prediction* of business failure from (usually publicly available) financial company data [9,24]. This approach is not primarily concerned with finding causes of undesirable behaviour, since the main purpose is prediction. Such an assessent may be very useful for potential investors and creditors, but is of limited value as a diagnostic tool for the company itself.

7.3. Other diagnostic tools

The approach discussed in this paper is primarily directed at a formalisation of the diagnostic process to the level that it may be implemented in computer algorithms to support human decision makers. There are a number of other techniques to support the diagnostic process more directed at model construction and visual representation of the model. A well-known example of such a technique in quality management are cause-and-effect (fishbone) diagrams [20]. Other techniques that may be helpful in model construction, e.g. the theory of Personal Constructs, can be found in the literature on knowledge elicitation [14]. The important problem of model construction and knowledge elicitation is however beyond the scope of this contribution.

Part II: Application

8. The interfirm comparison model

The model that we present in this section has been taken from the book "Interfirm Comparison" [19]. The model was originally developed for a case study, in order to demonstrate how management can use interfirm comparison (IFC) to diagnose the strengths and weaknesses of its business, and take remedial action. This case study involves the comparison of nine firms that operate in a section of the mechanical engineering industry. The firms in this particular section manufacture a class of industrial products needed by their customers as components of their own products.

Translation of the IFC model into the formalism for business models that we presented in Section 4, yields the following equations and qualitative relations:

1. $\mathrm{ROA} = \mathrm{PM} \times \mathrm{TA}$,
2. $\mathrm{PM} = 1 - (\mathrm{AC} + \mathrm{PC} + \mathrm{RDC} + \mathrm{DMC})$,
3. $\mathrm{TA} = 1 \div \mathrm{OA}$,
4. $\mathrm{OA} = \mathrm{CA} + \mathrm{FA}$,
5. $\mathrm{PC} = \mathrm{MC} + \mathrm{WLC} + \mathrm{WS} + \mathrm{OPC}$,
6. $\mathrm{CA} = \mathrm{FGS} + \mathrm{DEB} + \mathrm{MS} + \mathrm{WIP}$,
7. $\mathrm{FA} = \mathrm{LB} + \mathrm{PAM} + \mathrm{OFA}$,
8. $\mathrm{WIP} \leftarrow \{(\mathrm{SPP}, \mathrm{neg})\}$,
9. $\mathrm{MC} \leftarrow \{(\mathrm{MU}, \mathrm{pos}), (\mathrm{BC}, \mathrm{pos})\}$,
10. $\mathrm{DMC} \leftarrow \{(\mathrm{SPP}, \mathrm{pos}), (\mathrm{EX}, \mathrm{pos})\}$,
11. $\mathrm{FGS} \leftarrow \{(\mathrm{SPP}, \mathrm{pos})\}$,
12. $\mathrm{DEB} \leftarrow \{(\mathrm{EX}, \mathrm{pos})\}$.

Note that the quantitative part of the model (expressions 1–7) consists of five additive expressions (2 and 4–7), and two non-additive expressions (1 and 3) that are monotonic on $\mathbb{R}^+$. Expressions 8–12 represent the relevant causal knowledge of the problem domain.

We first explain the different variables and relations that appear in the model. ROA gives an overall indication of how profitably a company's management is using the resources available to it. This variable is split up into profit margin (PM) and turnover of assets (TA) (expression 1). Profit margin shows what profit has been earned on sales and turnover of assets shows how intensively the firm uses the available assets. The firm's profit margin on sales is determined by its departmental cost ratios (administrative costs (AC), production costs (PC), research & development costs (RDC) and distribution and marketing costs (DMC)), which express the cost falling under four headings as a percentage of sales (expression 2). The higher the cost ratios, the lower will be the profit margin on sales. The variables work sub-contracted (WS), materials costs (MC), works labour costs (WLC) and other production costs (OPC) provide a breakdown of PC (expression 5). IFC of these ratios, which express the PC incurred under four headings as fractions of the sales value of goods produced, will show how far each of the main areas of PC accounts for differences between a firm's PC to sales and those of others.

A firm whose cost of bought-out components is high in relation to its total material cost, will tend to have a relatively high materials cost ratio (MC), because the cost of these parts will include the production and other costs incurred and the profit margin taken by the suppliers (relation 9). The materials-cost ratio is also affected by the kind of metal, namely ferrous or non-ferrous, that the company uses in its production process (MU). Non-ferrous metals are generally more expensive. The firm's distribution and marketing cost ratio (DMC) is mainly influenced by its fraction of export sales (EX) and its sales/production policy (SPP) (expression 10). A high fraction of export sales tends to lead to higher distribution and marketing costs. Also a firm that uses sales/production policy type *A*, i.e. predominantly quantity production of standard products made for stock in anticipation of orders, will generally have higher distribution and marketing costs than firms that use policy *B* or *C* (see Table 3).

TA is a simple transformation of operating assets (OA) (expressed as a fraction of sales), and indicates the asset utilization rate of a firm (expression 3). The lower the figure the higher will be the rate of asset utilization. OA itself is sub-divided into the utilization rate of current assets (CA) and fixed assets (FA) (expression 4). The variables work in progress (WIP), finished goods stock (FGS), materials stock (MS) and debtors (DEB) provide a breakdown of the current assets ratio (expression 6). Interfirm comparison of WIP, FGS, MS and DEB shows to what extent each of the main components of current assets accounts for interfirm differences in CA. The variables land and buildings (LB), plant and machinery (PAM) and other fixed assets (OFA) give a breakdown of the fixed-assets ratio (expression 7). The ratio of debt to sales (DEB) is influenced by the fraction of export sales since longer credits may have to be given to foreign customers (expression 12).

The graph in Fig. 1 shows the explanatory graph $E(\mathrm{IFC})$, as defined in Section 4, for the IFC model.

Most of the variables take their values from the real numbers, except for SPP and MU (metal used). Note that these variables appear only in qualitative relations. SPP can have the value *A*, *B* or *C* (see Table 3), which means that it is measured on a nominal scale. With respect to the relations between SPP and other variables, this scale is converted to an ordinal scale, i.e. the basic operation on its elements is determination of greater or less. In this model, the values *A*, *B* and *C* are ordered: $A > B = C$. The variable MU stands for the metal that is predominantly used in the production process. This variable can take on two values: predominantly ferrous and predominantly non-ferrous. Regarding the relation between MU and materials cost/sales value of production (MC) the values are ordered non-ferrous > ferrous, reflecting the fact that non-ferrous metals are more expensive.

Table 3
Sales and production policy

Type	Sales/production policy
A	Predominantly quantity production of standard products made for stock in anticipation of order
B	Predominantly quantity production to customers' requirements after receipt of orders
C	Predominantly small quantity production to customers' orders, with occasional small batch production for stock

Table 4 gives the data of nine different firms operating in the mechanical engineering industry, for all variables that appear in the IFC model. The data of these firms have been taken from [19], with the additional note however that the values of the variables SPP (*A*, *B* and *C*) and MU (ferrous and non-ferrous), have been replaced by "order preserving" integer values, in order to enable the computation of an industry average. The column "average" has been added to the original data to represent norm behaviour in the examples that follow. The average values have been determined as follows:

- For all variables that do not appear on the left-hand side of an equation, the average has been computed directly by taking the mean value of all nine firms.
- For each equation $y = f(\mathbf{x})$, $\bar{y}$ has been computed by taking $\bar{y} = f(\bar{\mathbf{x}})$.

We follow this procedure in order to guarantee that the "average state" is internally consistent, which would not be the case if all averages would have been computed directly by taking the mean value of the nine firms.

Apart from being an illustration of the theory, the example explanations we examine will at the same time be a *test* of the theory, since we will compare the explanations provided in the book from which this case was taken [19] to the results obtained by our theory.

In [19] (p. 35) the following explanation is given for firm 1 relatively, i.e. compared to other firms, high ROA:

> *Why is firm 1's return on operating assets relatively high? Comparisons of ratios 2 (PM) and 3 (TA) with those of the other firms show that the firm's high ratio 1 (ROA) is due to a*

Table 4
Data for IFC

Variables	Firms									
	1	2	3	4	5	6	7	8	9	Av
ROA	0.251	0.239	0.189	0.161	0.133	0.132	0.088	0.079	0.041	0.137
PM	0.19	0.199	0.151	0.099	0.103	0.115	0.087	0.089	0.047	0.12
TA	1.32	1.2	1.25	1.63	1.29	1.15	1.01	0.9	0.87	1.14
OA	0.758	0.833	0.8	0.613	0.775	0.869	0.99	1.111	1.149	0.877
PC	0.628	0.635	0.711	0.747	0.725	0.719	0.754	0.774	0.802	0.722
RDC	0.005	0.01	0.009	0.007	0.007	0.011	0.014	0.0	0.002	0.007
DMC	0.109	0.116	0.047	0.072	0.062	0.058	0.073	0.046	0.064	0.072
AC	0.068	0.04	0.082	0.075	0.103	0.097	0.072	0.091	0.085	0.079
MC	0.32	0.287	0.339	0.301	0.397	0.316	0.336	0.347	0.358	0.333
WLC	0.165	0.221	0.232	0.283	0.157	0.241	0.248	0.274	0.289	0.234
OPC	0.143	0.127	0.14	0.092	0.104	0.113	0.114	0.153	0.155	0.127
WS	0.0	0.0	0.0	0.071	0.067	0.049	0.056	0.0	0.0	0.027
CA	0.465	0.481	0.412	0.35	0.369	0.449	0.549	0.582	0.608	0.474
FA	0.293	0.352	0.388	0.263	0.406	0.42	0.441	0.529	0.541	0.404
MS	0.08	0.11	0.081	0.079	0.068	0.092	0.1	0.101	0.082	0.088
WIP	0.043	0.04	0.063	0.062	0.083	0.106	0.188	0.225	0.245	0.117
FGS	0.132	0.102	0.047	0.037	0.039	0.045	0.044	0.053	0.057	0.062
DEB	0.21	0.229	0.221	0.172	0.179	0.206	0.217	0.203	0.224	0.207
LB	0.13	0.158	0.194	0.169	0.208	0.214	0.311	0.277	0.268	0.214
PAM	0.16	0.189	0.19	0.089	0.194	0.203	0.123	0.246	0.264	0.184
OFA	0.003	0.005	0.004	0.005	0.004	0.003	0.007	0.006	0.009	0.005
EX	0.16	0.4	0.15	0.0	0.25	0.23	0.1	0.12	0.3	0.19
SPP	2	2	1	1	1	1	1	1	1	1.22
BC	0.32	0.39	0.29	0.33	0.56	0.35	0.4	0.37	0.43	0.382
MU	2	1	2	1	2	1	1	2	1	1.44

combination of a comparatively high profit on sales and a comparatively fast turnover of assets.

The event to be explained is $\langle$firm 1, ∂ROA = high, industry average$\rangle$. In the quoted text, industry average is not mentioned explicitly as a reference object. It is stated however that the values of PM and TA for firm 1, are compared with those of the other firms. Therefore it is justified to take the industry average as the reference object, in order to make a comparison between the textbook analysis, and the results of our model of explanation. The model yields the following results, taking $T^+ = T^- = 0.85$. In Table 5 comparison is made between ROA of firm 1 and industry average (norm). From the data in Table 5 it follows that $C_p^+ = \{\mathrm{PM}, \mathrm{TA}\}$, since both PM and TA contributed to the difference between norm value and actual value, and both are needed to explain the desired fraction of $\inf(C^+, \mathrm{ROA})$. Obviously, $C_p^- = \{\ \}$.

Table 5
Data for explanation of ∂ROA = high

Firm 1	Norm	Actual	inf
ROA	0.137	0.251	
PM	0.12	0.19	0.0796
TA	1.14	1.32	0.0214

Comparison of human analysis and the result of our explanation method shows some noticable similarities. Firstly, both the human analyst and the model explain the relatively high ROA in terms of its right-hand side variables, profit on sales (PM) and TA. The model consistently uses the industry average as a comparison to explain the relatively high ROA of firm 1. The second sentence of the textbook analysis gives no clue about how exactly these "comparisons of ratios 2 and 3 with those of the other firms" are made. Both the human analyst and the model state that both PM and TA had a significant influence.

The analysis continues as follows ([19], p. 35):

Why is the firm's profit on sales ratio relatively high ? The most immediate answer is provided by its cost/sales ratios ... Firm 1's profit on sales is high mainly because its production cost ratio is the lowest of all ... On the other hand, its distribution and marketing cost ratio is almost the highest of all.

Analogous to the previous example, the event to be explained is specified as $\langle$firm 1, ∂PM = high, industry average$\rangle$. Table 6 summarizes the model results for the explanation of firm 1's relatively high profit on sales. From the data in Table 6 it follows that $C_p^+ = \{\text{PC}\}$ and $C_p^- = \{\text{DMC}\}$. Notice that neither the human analyst nor the model mention that firm 1's administrative cost/sales ratio is below average. The reason is that its contribution to the overall contributing influence ($\inf(C^+, \text{PM}) = 0.107$) on profit on sales is negligible. The same reasoning goes for research and development cost. This shows that both the human analyst and the model of explanation tend to leave insignificant influences out of the explanation. The human analyst also mentions a counteracting influence, by indicating that firm 1's marketing cost ratio is almost the highest of all firms considered. This counteracting influence is also noticed by the explanatory model, since $C_p^- = \{\text{DMC}\}$.

The foregoing examples both involved quantitative equations. The next example shows how qualitative relations are used to generate explanations. Again we make a comparison between textbook analyses and model results. In p. 36 of [19] we find the following analysis:

Table 6
Data for explanation of ∂PM = high

Firm 1	Norm	Actual	inf
PM	0.12	0.19	
AC	0.079	0.068	0.011
PC	0.722	0.628	0.094
RDC	0.007	0.005	0.002
DMC	0.072	0.109	−0.037

Reverting to firm 1, ... this firm applies sales/production policy A; it follows that its relatively high distribution and marketing cost ratio can probably be regarded as being part and parcel of its sales/production policy and therefore a condition of its success. The other factor ... – a high export percentage – would not have caused firm 1's distribution and marketing cost ratio to be high, since its percentage of export sales is comparatively small.

Again, industry average seems to be the proper reference object, i.e. the event to be explained is $\langle$firm 1, ∂DMC = high, industry average$\rangle$. From Table 7 it can be concluded that $C_p^+ = \{\text{SPP}\}$ and $C_p^- = \{\text{EX}\}$. The method of explanation generation for qualitative relations is clearly reflected in the above textbook fragment. The percentage of export sales of firm 1 cannot explain its relatively high distribution and marketing costs because its export percentage is actually *below* average and there is a positive relation between export sales and distribution and marketing costs. This is reflected by the fact that the model of explanation lists percentage of export sales (EX) as a counteracting influence. An explanation can be found however by looking at firm 1's SPP. This is exactly the explanation that the model gives, since SPP is listed as the only contributing influence.

The previous examples of different one-level explanations, show that the results of our method correspond to textbook analyses. We now provide an example of a complete diagnosis. The sole performance indicator in the IFC model is return on assets (ROA). We assume the norm model specifies that ROA should not be more than 50% below industry average.

The diagnosis will be performed for firm 9, which is doing particularly badly compared to industry average. Problem identification yields the

Table 7
Data for explanation of ∂DMC = high

Firm 1	Norm	Actual	∂	sign
DMC	0.072	0.109	High	
SPP	1.22	2	High	pos
EX	0.19	0.16	Low	pos

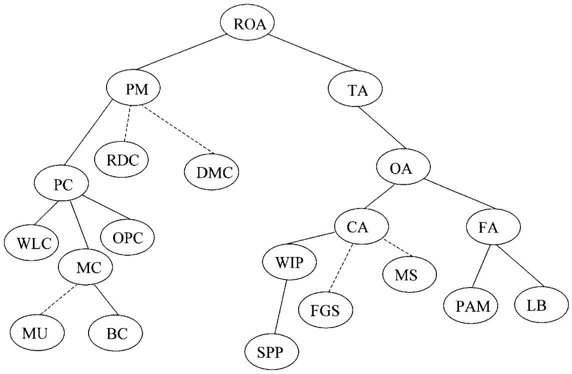

Fig. 2. Diagnosis for $S = \{\partial \text{ROA} = \text{low}\}$ at firm 9.

set of symptoms $S = \{\partial \text{ROA} = \text{low}\}$ since the relative difference between norm value and actual value, $(0.041 - 0.137)/0.137 = -0.7$, i.e. 70% below industry average.

The diagnosis starts with a one level explanation of $\partial \text{ROA} = \text{low}$. Application of Definition 1 yields: $C_p^+ = \{\text{PM}, \text{TA}\}$ and $C_p^-(\text{ROA}) = \{\ \}$. Fig. 2 summarizes the results of the complete diagnostic process, where dashed lines indicate counteracting causes. Since there is only one symptom to be explained, the diagnosis contains one maximal explanation. Thus, Fig. 2 actually depicts the maximal explanation, as specified in Definition 3, for $\partial \text{ROA} = \text{low}$.

9. Conclusions

In normative research on managerial decision making, the importance of problem diagnosis is frequently stressed, yet a formal description of how diagnosis is to be performed is hard to find. In this paper, we presented a formal framework for explanation and diagnosis of company performance with both qualitative and quantitative information. For the construction of explanations, the canonical format of aleatory explanations is adapted to the requirements of the business domain. The sets of contributing and counteracting causes are reduced to *parsimonious* sets, in order to avoid the inclusion of insignificant causes. For the determination of contributing and counteracting causes, we developed the *inf-measure* which embodies a kind of *ceteris paribus* reasoning.

We illustrated the applicability of our method by comparison of its results with textbook analyses. This comparison showed that there is a large correspondence. Limited availability of human analyses precluded a more elaborate validation however. More elaborate validation experiments are clearly required in future research.

The results of research in this area can be used to develop analytical tools for management or other professionals concerned with the assessment of business performance. We used the diagnostic method as part of a financial diagnosis system to be used by investment analysts. The research presented in this paper is a step in that direction, yet there remain many interesting questions to be solved. One issue we would like to mention especially is the link between diagnosis and therapy. The primary cause of a symptom may, for example, be a variable over which management has no control. Hence therapeutic action to remove the undesired situation may have to be directed at other influencing variables, and the link between diagnosis and therapy is not as straightforward as one might wish. This question will also be addressed in future research.

References

[1] N.H.A. Mohammed, J.F. Courtney, D.B. Paradice, A prototype DSS for structuring and diagnosing managerial problems, IEEE Transactions on Systems, Man and Cybernetics 18 (6) (1988) 899–907.

[2] R. Berndsen, H. Daniels, Casual reasoning and explanation in dynamic economic systems, Journal of Economic Dynamics and Control 18 (1) (1994) 251–271.

[3] M.J. Bouwman, A.P. Frishkoff, P. Frishkoff, How do financial analysts make decisions? A process model of the investment screening decision, Accounting Organizations and Society 12 (1) (1987) 1–29.

[4] M.J. Bouwman, Human diagnostic reasoning by computer: An illustration from financial analysis, Management Science 29 (6) (1983) 653–672.

[5] C.M. Brugha, The structure of qualitative decision making, European Journal of Operational Research 104 (1) (1998) 46–62.

[6] A. Charnes, W. Cooper, A.Y. Lewin, L.M. Seiford, Data Envelopment Analysis: Theory, Methodology and Applications, Kluwer Academic Publishers, Norwell, MA, 1994.

[7] A. Charnes, W.W. Cooper, E.L. Rhodes, Measuring the efficiency of decision making units, European Journal of Operational Research 2 (6) (1978) 429–444.
[8] J.F. Courtney, D.B. Paradice, N.H.A. Mohammed, A knowledge-based DSS for managerial problem diagnosis, Decision Sciences 18 (1987) 373–399.
[9] A.I. Dimitras, S.H. Zanakis, C. Zopounidis, A survey of business failures with an emphasis on prediction methods and industrial applications, European Journal of Operational Research 90 (3) (1996) 487–513.
[10] C. Eden, S. Jones, D. Sims, Thinking in Organizations, MacMillan, London, 1979.
[11] C. Eden, S. Jones, D. Sims, Messing About in Problems, Pergamon Press, Oxford, 1983.
[12] A.J. Feelders, A model of explanation for financial knowledge-based systems, Computer Science in Economics and Management 5 (1992) 119–132.
[13] A.J. Feelders, Diagnostic reasoning and explanation in financial models of the firm, Ph.D. Thesis, Tilburg University, 1993.
[14] M. Firlej, D. Hellens, Knowledge Elicitation: A Practical Handbook, Prentice-Hall, Hemel Hempstead, 1991.
[15] F. Harary, R.Z. Norman, D. Cartwright, Structural Models: An Introduction to the Theory of Directed Graphs, Wiley, New York, 1965.
[16] G. Hesslow, Explaining differences and weighting causes, Theoria 49 (1984) 87–111.
[17] P.W. Humphreys, The Chances of Explanation, Princeton University Press, Princeton, NJ, 1989.
[18] A. Huxor, Explanation and understanding: An architecture for economic elaboration. Technical Report TR-LP-26, ECRC, München, Germany, 1988.
[19] H. Ingham, L.T. Harrington, Interfirm Comparison, Heinemann, London, 1980.
[20] K. Ishikawa, Guide to Quality Control, Asian Productivity Organization, Tokyo, 1976.
[21] R.S. Kaplan, D.P. Norton, The balanced scorecard – measures that drive performance, Harvard Business Review, January/February 1992, pp. 71–79.
[22] D.W. Kosy, Applications of explanation in financial modeling, in: L.E. Widman, K.A. Loparo, N.R. Nielson, (Eds.), Artificial Intelligence, Simulation and Modeling, Wiley, New York, 1989, pp. 487–509.
[23] D.W. Kosy, B.P. Wise, Self-explanatory financial planning models, in: Proceedings of AAAI-84, Los Altos, CA, Morgan Kaufmann, Los Altos, CA, 1984, pp. 176–181.
[24] P.P.M. Pompe, A.J. Feelders, Using machine learning neural networks and statistics to predict corporate bankruptcy: A comparative study, in: P. Ein-Dor (Ed.), Artificial Intelligence in Economics and Management, Kluwer Academic Publishers, Boston, MA, 1996, pp. 3–19.
[25] D.S. Weld, J. deKleer, Readings in Qualitative Reasoning about Physical Systems, Morgan Kaufmann, San Mateo, CA, 1990.