Fraunhofer Institut
Techno- und
Wirtschaftsmathematik

P. Lang, A. Sarishvili, A. Wirsen

Blocked neural networks for knowledge extraction in the software development process

Berichte des Fraunhofer ITWM, Nr. 56 (2003)

ISSN 1434-9973

Bericht 56 (2003)

Die Veröffentlichungen in der Berichtsreihe des Fraunhofer ITWM können bezogen werden über:

Fraunhofer-Institut für Techno- und
Wirtschaftsmathematik ITWM
Gottlieb-Daimler-Straße, Geb. 49

67663 Kaiserslautern
Germany

Telefon: +49 (0) 6 31/2 05-32 42
Telefax: +49 (0) 6 31/2 05-41 39
E-Mail: info@itwm.fraunhofer.de
Internet: www.itwm.fraunhofer.de

Vorwort

Das Tätigkeitsfeld des Fraunhofer Instituts für Techno- und Wirtschaftsmathematik ITWM umfasst anwendungsnahe Grundlagenforschung, angewandte Forschung sowie Beratung und kundenspezifische Lösungen auf allen Gebieten, die für Techno- und Wirtschaftsmathematik bedeutsam sind.

In der Reihe »Berichte des Fraunhofer ITWM« soll die Arbeit des Instituts kontinuierlich einer interessierten Öffentlichkeit in Industrie, Wirtschaft und Wissenschaft vorgestellt werden. Durch die enge Verzahnung mit dem Fachbereich Mathematik der Universität Kaiserslautern sowie durch zahlreiche Kooperationen mit internationalen Institutionen und Hochschulen in den Bereichen Ausbildung und Forschung ist ein großes Potenzial für Forschungsberichte vorhanden. In die Berichtreihe sollen sowohl hervorragende Diplom- und Projektarbeiten und Dissertationen als auch Forschungsberichte der Institutsmitarbeiter und Institutsgäste zu aktuellen Fragen der Techno- und Wirtschaftsmathematik aufgenommen werden.

Darüberhinaus bietet die Reihe ein Forum für die Berichterstattung über die zahlreichen Kooperationsprojekte des Instituts mit Partnern aus Industrie und Wirtschaft.

Berichterstattung heißt hier Dokumentation darüber, wie aktuelle Ergebnisse aus mathematischer Forschungs- und Entwicklungsarbeit in industrielle Anwendungen und Softwareprodukte transferiert werden, und wie umgekehrt Probleme der Praxis neue interessante mathematische Fragestellungen generieren.

Prof. Dr. Dieter Prätzel-Wolters
Institutsleiter

Kaiserslautern, im Juni 2001

Blocked neural networks for knowledge extraction in the software development process

Patrick Lang, Alex Sarishvili, Andreas Wirsen

October 2003

Abstract

One of the main goals of an organization developing software is to increase the quality of the software while at the same time to decrease the costs and the duration of the development process. To achieve this, various decisions effecting this goal before and during the development process have to be made by the managers. One appropriate tool for decision support are simulation models of the software life cycle, which also help to understand the dynamics of the software development process. Building up a simulation model requires a mathematical description of the interactions between different objects involved in the development process. Based on experimental data, techniques from the field of knowledge discovery can be used to quantify these interactions and to generate new process knowledge based on the analysis of the determined relationships. In this paper blocked neuronal networks and related relevance measures will be presented as an appropriate tool for quantification and validation of qualitatively known dependencies in the software development process.

Keywords: Blocked Neural Networks, Nonlinear Regression, Knowledge Extraction, Code Inspection

1 Introduction

During the last years various simulation tools focusing on different aspects in the software development process were presented in the literature. They roughly can be classified into continuous (system dynamics), discrete event or hybrid simulation models. We especially are involved in developing a discrete event model, which focuses mainly on the inspection process in the coding phase and the test phase of the software life cycle (see [6],[7]). In comparison to a continuous simulation model, choosing a discrete event approach allows a more detailed representation of the organizational issues, products and resources. It is i.e. possible to model programmers or inspectors with different skills or to consider different types of software items with varying complexity and size.
Building up a simulation model requires the determination of input output relationships at the different sub-processes of the software development process,

which are in case of considering a waterfall model the requirement, design, coding and test phase. In software engineering the software development process is described by static qualitative models like control diagrams, flow diagrams and cause-effect diagrams. Figure 1 shows as an example the control diagram, the flow diagram and cause-effect diagram of an inspection process. These models are in contrast to a simulation model not appropriate to study the dynamic effects occurring during the software development process. However, they provide a general understanding concerning the chronology of tasks (control diagram), the flow of the objects (flow diagram) and the qualitative dependencies of objects (cause-effect diagrams) and thus include the basic information needed for building up a simulation model.

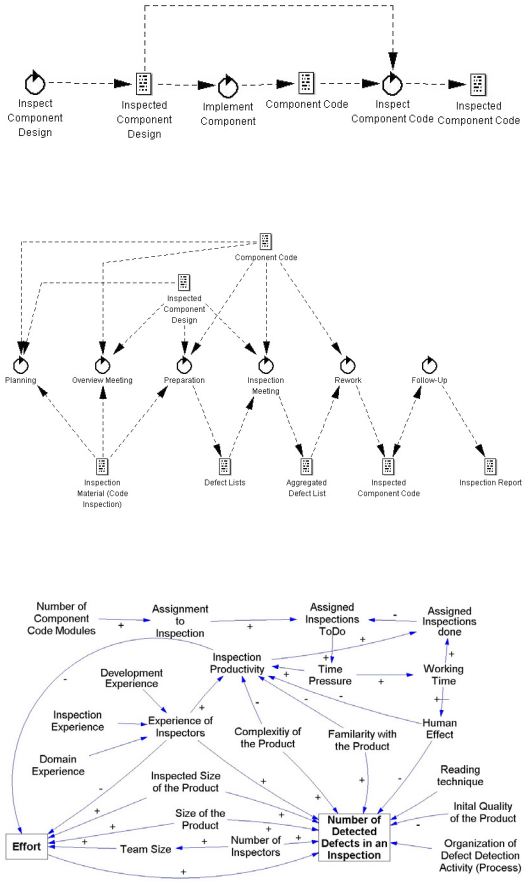

Figure 1: Qualitative models of the software inspection process. Top: Control Diagram. Middle: Flow Diagram. Bottom: Cause-Effect Diagram

In [6] the general method for the generation of our discrete event simula-

tion model is explained. The model can be built by following step by step the control and flow diagram (see figure 1). In this way activity blocks inside the simulation model with their related inputs and outputs, i.e. items, stuff, etc., are determined. Inside an activity block the relationships between certain variables, qualitatively described by the corresponding cause-effect diagrams, then have to be quantified by mathematical equations or logical rules. A cause-effect diagram generally distinguishes between three types of variables:

- process input variables, which do not depend on other variables,
- process output variables, which do not effect other variables and
- internal process variables, which are explaining other variables and are also explained by other variables.

Based on the cause-effect diagram one step by step chooses each of the internal and output variables as the explained variable and their corresponding predecessors as explaining variables. The related input output relationships then have to be determined as a logical rule or a mathematical function.

Possible methods for quantifying the qualitatively known input output relationships are expert interviews, pragmatic models, stochastic analysis and knowledge discovery techniques. The choice of the actual technique depends of course on the data or information available, i.e. measurement data from experiments, linguistic descriptions, etc..

In this paper we assume that sufficient measurement data is given such that an application of knowledge discovery techniques is possible. In a first step these methods will be used for the quantification of the input output relationships. Then, based on the identified mappings, new insight and rules for parts of the considered process will be generated. In the example at the end we investigate mathematical descriptions determining the *number of major defects found in an inspection process* depending on the *effort*, the *size of the products* and other variables as can be seen in the related cause-effect diagram (see figure 1).

The simplest approach for the quantification of an input output relationship in form of a mathematical equation is to consider a linear regression problem:

$$x_j = \sum_{i=1}^{d} Y_{ji} a_i + \epsilon_j, \;\; j = 1, \ldots, n,$$

where $x_j \in \mathbb{R}$ contains the j-th measurement of the output variable, $Y_j \in \mathbb{R}^d$ are vectors containing the measurements of the related input variables, $a_i \in \mathbb{R}$ denote the d unknown regression coefficients and ϵ_j is the measurement error of the j-th variable.

However, due to saturation effects the influence of some of the input variables, like *skills of programmers or inspectors* or other human and document properties, obviously is nonlinear. Therefore it seems quite reasonable to use a generalized regression model.

More especially, we will consider an additive nonlinear regression model (AN) of the form

$$x_j = \sum_i a_i m_i(Y_{ji}) + \epsilon_j,\ j = 1, \ldots, n,$$

where $m_i : \mathbb{R} \to \mathbb{R},\ i = 1, \ldots, d$ are nonlinear twice differentiable functions depending on further regression coefficients and $a_i \in \mathbb{R}$ are again regression coefficients. AN models are reasonable generalizations of classical linear models, since they conserve the interpretability property of linear models and simultaneously are able to reproduce certain nonlinearities in the data. It also is possible to calculate and interpret the partial derivatives of an AN model. The importance of the partial derivatives lies in relevance and sensitivity analysis.

The additive nonlinear regression function can be approximated by different methods. The most common of them are so called back fitting algorithms [8] with various smoothing operators, like:

- Univariate regression smoothers such as local polynomial regression.
- Linear regression operators yielding polynomial or parametric spline fits.
- More complex operators such as surface smoothers.

In this paper we consider specially structured (block-wise) neural networks for the estimation of the AN model. In [4], [9] it was shown that fully connected neural networks are able to approximate arbitrary continuous functions with arbitrary accuracy, furthermore in [10] it was proven that neural networks are able to approximate the derivatives of regression functions. This result was assigned in [16] to block-wise neural networks as estimators for nonlinear additive twice continuously differentiable regression functions and their derivatives. The network function consists of input and output weights for each unit in the hidden layer, which have to be estimated from given measurement data. This is done by minimizing the mean squared error over a training set. The performance of the network is measured by the prediction mean squared error, which is estimated by cross validation [3].
The network function estimated by the neural network will be used when the model is created, i.e. it will be implemented in the model at the considered node of the cause-effect diagram to calculate the output for a given input during the simulation is running. Here one has to mention, that the structure of the network in general remains unchanged for the developed simulation model since it is based on qualitative models while the weights of the network function might change for different software development processes and thus the neural network has to be retrained with respect to the considered process. In the simulation model the values of input variables of the equations have to be provided in the considered activity block.

Now it will be shown that relevance measures calculated on the identified network function also are an important tool in the context of the software development process. The problem, which often occurs especially in modelling

the software development process is, that not for all input variables measurement data are available. The granularity of the model determines the minimal amount of measurement data needed for rule generation since all input and output variables of the underlying qualitative model should be used at this step. In case of missing measurement data for one or more variables one has to make further assumptions or skip these variables. Relevance measures in this case help to determine the impact of each input variable with respect to the output variable. By considering the validation results and the corresponding relevance measure, one can easily verify whether the estimated functional dependencies describe the input output relationship in a sufficient manner, if the impact of a skipped variable is too large or if an explaining (input) variable is missing in the underlying qualitative model. In the latter case, the missing variable should be determined by analyzing the set of all measured variables e.g. by a case-based-reasoning method (see [20]) and a new rule has to be generated for by incorporating the identified missing input variables or by applying other knowledge discovery techniques. If a variable has only small relevance over its whole measurement range it is redundant, i.e. it does not explain the considered output and thus can be skipped. Thus, a relevance measure can also be used to validate the qualitative description of the dependencies given in the cause-effect diagrams since the inputs for a node in the diagram are assumed to be not redundant. Also the size of the impact of every input variable for a given data set is available and might give the manager a new insight into the considered software development process. Thus, the relevance measure might help to provide the manager a new rule of thumb.

Using the blocked neural network approach two aspects mentioned above will be considered, which are

1. the quantification of the qualitative models
2. and analyzing of the determined mathematical equations in order to find a more deep insight into the input output relationship by relevance or sensitivity analysis.

Currently, we analyze data on historical software development processes coming from a large company. Unfortunately, these data (which were not collected for the purpose of fitting a simulation model) cover only some of the variables required for building the desired discrete event simulation model of the inspection process. For instance, considering the cause-effect diagram of the inspection process (see figure 1), information on the assignment of tasks to persons, skills and individual working times is missing. However, our approaches from neural networks describe an appropriate idea how the input output relationship could be achieved and how to determine the impact of the input variables with respect to the considered output based on the identified mapping.

2 Neural network modelling

Neural networks provide a convenient language game for nonlinear modelling. They are typically used in pattern recognition, where a collection of features is presented to the network, and the task is to assign the input to one or more classes. In [13] it is shown that arbitrary complex decision regions, including concave regions, can be formed using four-layer neural networks. In [11] the ability of three-layer networks to form several complex decision regions in pattern recognition applications is demonstrated by simulations.

Another typical use for neural networks is the field of nonlinear regression problems. In [2], [4], [9], [1], it is shown independently, that neural networks with linear output and single hidden layer with sigmoid activation function can approximate any continuous regression function uniformly on compact domains.

Neural networks typically exhibit two types of behavior. If no feedback loops are present in the network connections, the signal produced by an external input moves in only one direction and the output of the network is given by the output of the last layer neurons. In this case a neural network behaves mathematically like a static nonlinear mapping of the inputs. This feedforward type of networks therefore is most often used for nonlinear function approximation. The second kind of network behavior is observed when feedback loops are present. In this case the network behaves like a dynamical system, and the outputs of the neurons are functions of time. The neuron outputs can for example oscillate, or converge into a steady state.

In this paper we consider feedforward neural networks with one hidden layer to solve regression problems occurring during modelling the software life cycle. The structure of such a neural network is depicted in figure 2.

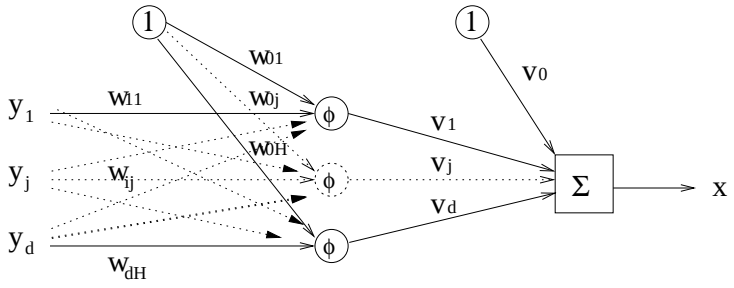

Figure 2: Feedforward neural network.

The output of the feedforward neural network is computed as:

$$f_{nn}(Y,\theta) = v_0 + \sum_{h=1}^{H} v_h \phi\big(Y \cdot w_h\big), \tag{1}$$

where $\theta = (w'_1, ..., w'_H, v_0, ..., v_H)$ is the vector of network weights, with $w'_h = (w_{0h}, ..., w_{dh})$ for $h = 1, ..., H$ and H denotes the number of neurons in the

hidden layer. Moreover, ϕ is the activation function from each hidden unit and $Y = (1, y_1, ..., y_d)$ denotes the vector of inputs. More about the special architecture of the neural networks considered in this paper will be discussed in the next sections.

Based on the given input output data the network parameters are estimated via convenient learning algorithms. In the case of feed forward networks usually a variant of the famous backpropagation algorithm is used.

2.1 Variable significance testing

If one not only is interested in obtaining a good input output approximation, but one further wants to get more insight into the structure of the underlying function, different postprocessing methods can be applied to the trained network. In order to derive the significance of each neural network input variable based on the derived function various statistical hypothesis tests can be used. In general, this involves the following main steps:

- the definition of a convenient relevance measure,
- estimating the defined relevance of each input variable with respect to the corresponding model output,
- estimating the sampling variability of the selected relevance measure,
- testing the null hypothesis of irrelevance.

In the next section we introduce some classical relevance measures that are often used in the case of neural networks.

2.2 Classical relevance measures in the neural network approach

We consider the following regression problem:

$$x_i = M(y_{1i}, ..., y_{di}) + \epsilon_i, \quad i \in [1, ..., N],$$

where ϵ_i is i.i. $N(0, \sigma^2)$ distributed noise, the function $M : \mathbb{R}^d \to \mathbb{R}$ is Borel measurable and twice differentiable at any $i \in [1, ..., N]$. We now approximate the true regression function M with a neural network approximator:

$$\begin{aligned} \hat{x}_i &= f_{nn}(y_{1i}, ..., y_{di}, \hat{\theta}) + \epsilon_i, \\ \hat{\theta} &= arg \min_{\Theta \in \Theta_H} \left(\sum_{i=1}^{N} (x_i - f_{nn}(y_{1i}, ..., y_{di}, \Theta))^2 \right), \end{aligned}$$

where Θ_H is a compact subset of the parameter (weight) space.

The most common measure of relevance is the average derivative (AD), since the average change in $\hat{x}$ for a very small perturbation $\delta y_j \longrightarrow 0$ in the independent variables y_j is simply given by:

$$AD(y_j) = \frac{1}{N} \sum_{i=1}^{N} \frac{\partial \hat{x}_i}{\partial y_{ji}}, \tag{2}$$

where $\hat{x}$ is the vector of estimated regression outputs. Sometimes $\hat{x}$ is sensitive to y_j only for a small percentage of input vectors y_j. Such requirements give rise to the following measure of relevance, that can be quite important in the context of particular applications.

$$MaxD(y_j) = \max_{i=1,...,N} \left| \frac{\partial \hat{x}_i}{\partial y_{ji}} \right|.$$

Another quantification of $\hat{x}$'s sensitivity to y_j, is the average percentage change in $\hat{x}$ for a one percent change in y_j, a measure that is commonly known as the "Average Elasticity" of $\hat{x}$ to y_j [15]:

$$AvE(y_j) = \frac{1}{N} \sum_{i=1}^{N} \left(\left| \frac{\partial \hat{x}_i}{\partial y_{ji}} \right| \right) \left(\left| \frac{y_{ji}}{\hat{x}_i} \right| \right), \ \hat{x}_i \neq 0 \ \ \forall \, i = 1, ..., N. \tag{3}$$

Another measure describes the average contribution of the y_j's to the magnitude of the gradient vector. Therefore a measure of sensitivity, namely the "standard deviation" of the derivatives across the sample measuring the dispersion of the derivatives around their mean, is computed:

$$SD(y_j) = \left[\frac{1}{N} \sum_{i=1}^{N} \left(\left| \frac{\partial \hat{x}_i}{\partial y_{ji}} \right| - \frac{1}{N} \sum_{k=1}^{N} \left| \frac{\partial \hat{x}_k}{\partial y_{jk}} \right| \right)^2 \right]^{\frac{1}{2}}.$$

Normalizing $SD(y_j)$ by the mean, provides us with the coefficient of variation that is the standard deviation per unit of sensitivity:

$$CV(y_j) = \frac{SD(y_j)}{\frac{1}{N} \sum_{k=1}^{N} \left| \frac{\partial \hat{x}_k}{\partial y_{jk}} \right|}.$$

All the described relevance measures share the disadvantage that for standard (fully connected) feedforward networks, the partial derivatives in general depend on all input variables. Therefore the interpretation in terms of the relevance of a single input variable is difficult. One method to avoid this problem is to set all variables except the considered variable y_j to their mean values. This however leads to a significant loss of information. Another method consists in using blocked neural networks as universal function approximators, a method that is described in the next sections.

At the end of this paper we will compare and discuss the results of a relevance analysis for a software development process based on the measures AD and AvE.

2.3 Additive nonlinear(AN) regression models

As described before, we use ANs for understanding software development processes profiting from the advantages of ANs, like interpretability and flexibility, compared to other methods. In this section we define a general form AN model and state main assumptions that should be fulfilled if ANs are chosen for modelling.

Let $Y \in \mathbb{R}^{d \times N}$ be a design data matrix, where each column refers to a single observation and each row to an attribute. In the following we describe an additive nonlinear regression problem.

Assumption 2.1. *The function $M : \mathbb{R}^d \longrightarrow \mathbb{R}$ that describes the true relationship between the dependent variables $x_i \in \mathbb{R}, \;\; i = 1, ..., N$, and the data design matrix Y exists.*

Assumption 2.2. *The conditional expectation function $M : \mathbb{R}^d \longrightarrow \mathbb{R}$,*

$$M(Y) = E\{x|(y_{1i}, ..., y_{di}) = Y\}$$

has an additive structure, i.e.

$$M(Y) = m_1(y_{1i}) + ... + m_d(y_{di}),$$

where $m_j : \mathbb{R} \longrightarrow \mathbb{R}, \;\; \forall j = 1, ..., d$.

Definition 2.3. *An Additive Nonlinear (AN(d)) model for any variable $x_i \in \mathbb{R}, \;\; i = 1, ..., N$ is defined by,*

$$x_i = m_1(y_{1i}) + ... + m_d(y_{di}) + \epsilon_i, \;\; i = 1, ..., N \tag{4}$$

where ϵ_i is i.i. $N(0, \sigma^2)$ distributed with finite variance.

In the course of this paper we estimate the function M by fitting feedforward neural networks with block structure to the data.

The optimal network hereby is determined by minimizing the mean squared prediction error.

2.4 AN model estimation with blocked neural networks

Taking into account the special structure of the composite function M, which results from summing up d functions of mutually different real variables, a feedforward network with one hidden layer and without "nonparallel" input to hidden layer connections seems to be convenient for its approximation. Figure 3

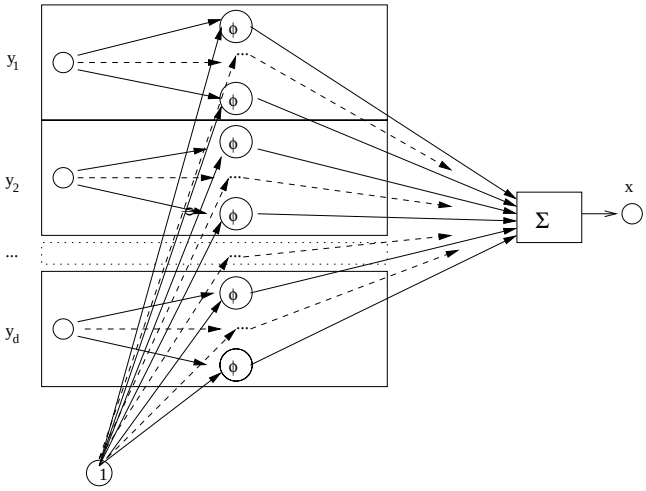

Figure 3: Blocked neural network.

shows such a blocked neural network with d inputs and one output. Especially, each neuron in the hidden layer accepts only one variable as input apart from the constant bias node.

The output of the blocked neural network is given as:

$$\begin{aligned} x_i &= f_{nn}^{(bl)}(y_{1i}, ..., y_{di}, \Theta) \\ &= \sum_{i=1}^{H(1)} v_i \phi(y_{1i} w_i + b_i) + ... + \sum_{i=H(d-1)+1}^{H(d)} v_i \phi(y_{di} w_i + b_i), \end{aligned} \tag{5}$$

where the difference $H(j) - H(j-1), \ \forall \ j = 2, ..., d$ is the number of neurons in block i, and $H(d)$ denotes the total number of neurons in the hidden layer. The w_js are the weights from the input to the hidden layer, the v_js are the weights from the hidden to the output layer, the b_js are the biases and Θ denotes the vector of all neural network parameters together. The neuron activation function is chosen to be of sigmoidal type, i.e.: $\phi(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$.

The neural network training consists of minimizing the mean squared error over all training samples resulting in the optimal network parameters:

$$\hat{\theta} = arg \min_{\Theta \in \Theta_H} \left(\sum_{i=1}^{N} (x_i - f_{nn}^{(bl)}(y_{1i}, ..., y_{di}, \Theta))^2 \right),$$

where Θ_H is a compact subset of the parameter space. In the following the supscripts of f are skipped.

The following theorem describes the approximation abilities of such a neural network.

Theorem 2.4. *Let $\phi(\cdot)$ be a nonconstant, bounded and monotonic increasing continuous function. Let K be a compact subset of $\mathbb{R}^d$ and $k \geq 1$ a fixed integer. Then any continuous mapping $F : K \longrightarrow \mathbb{R}$ with $F(y_1, ..., y_d) = f_1(y_1) + ... + f_d(y_d)$, where $f_j : \mathbb{R} \longrightarrow \mathbb{R}$, $j = 1, ..., d$ are continuous and $K \subset Domain(f_j)$, can be approximated in the sense of uniform topology on K by blocked networks with one hidden layer, where the hidden layer functions are chosen as $\phi(\cdot)$ and the input and output layer are defined by arbitrary linear functions.*

This theorem is identical to the general theorem presented by [4] except the assumption that the function F is a sum of d continuous functions and therefore continuous itself. Thus the proof is analogous to the proof in [4].

Note, that multilayer feedforward neural networks not only are capable of arbitrary accurate approximations for unknown mappings, but further they also can be used to estimate simultaneously the related derivatives, see [10].

2.4.1 Derivatives in blocked neural networks.

In this section we derive the important property of blocked neural networks consisting in the special relation between the expected partial derivatives of the network function with respect to the input variables and the partial derivatives with respect to the network weights. For simplification we consider a blocked neural network with only one neuron in each block, see figure 4. All results shown in this section are also transferable to arbitrary connected blocked neural networks.

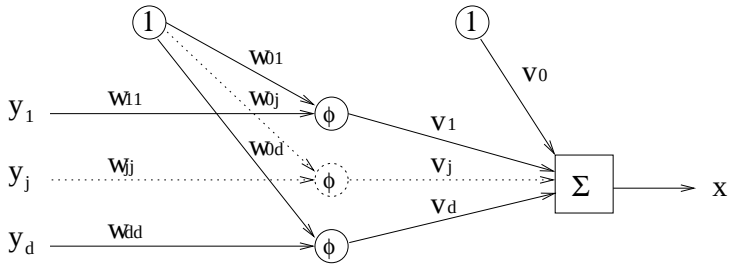

Figure 4: Blocked neural network with one neuron in each block.

The considered network function then has the form:

$$f(y, \theta) = v_0 + \sum_{j=1}^{d} v_j \phi(w_{0j} + w_{jj} y_j),$$

where y_j is the j-th coordinate of y. If ϕ is differentiable, as we always assume, the derivatives w.r.t. y_j and w.r.t. the parameter w_{jj} are given as:

$$\begin{aligned}\frac{\partial f(y,\theta)}{\partial y_j} &= v_j\phi'(w_{0j}+w_{jj}y_j)w_{jj}\\ \frac{\partial f(y,\theta)}{\partial w_{jj}} &= v_j\phi'(w_{0j}+w_{jj}y_j)y_j.\end{aligned}$$

Based on these identities the following theorem can be proven.

Theorem 2.5. *For a blocked neural network with one neuron in each block, we consider*

$$\begin{aligned}C_j &:= E\left(\left|\frac{\partial f(y,\theta)}{\partial w_{jj}}\right|\right),\\ R_{j,1} &:= E\left(\left|\frac{\partial f(y,\theta)}{\partial y_j}\right|\right),\\ R_{j,2}^2 &:= E\left(\left|\frac{\partial f(y,\theta)}{\partial y_j}\right|^2\right).\end{aligned}$$

a) If the j-th input variable is bounded with essential supremum $||y_j||_\infty$, then

$$|w_{jj}|C_j \le ||y_j||_\infty R_{j1}$$

b) In general, we have

$$|w_{jj}|C_j \le \sqrt{E(|y_j|^2)}R_{j2}.$$

Proof

a)

$$\begin{aligned}C_j &= E\left(\left|\frac{\partial f(y,\theta)}{\partial w_{jj}}\right|\right)\\ &= |v_j|E\big(|\phi'(w_{0j}+w_{jj}y_j)|\cdot|y_j|\big)\\ &\le |v_j|\cdot||y_j||_\infty E\big(|\phi'(w_{0j}+w_{jj}y_j)|\big)\\ &= \frac{||y_j||_\infty}{|w_{jj}|}E\left(\left|\frac{\partial f(y,\theta)}{\partial y_j}\right|\right)\\ &= \frac{||y_j||_\infty}{|w_{jj}|}R_{j,1}\end{aligned}$$

b)

$$\begin{aligned} C_j^2 &= \left(E\left|\frac{\partial f(y,\theta)}{\partial w_{jj}}\right|\right)^2 \\ &= v_j^2\left(E\left(|\phi'(w_{0j}+w_{jj}y_j)|\cdot|y_j|\right)\right)^2 \\ &\leq v_j^2 E(|\phi'(w_{0j}+w_{jj}y_j)|^2)E(|y_j|^2) \\ &= \frac{E(|y_j|^2)}{w_{jj}^2}E\left(\left|\frac{\partial f(y,\theta)}{\partial y_j}\right|^2\right) \\ &= \frac{E(|y_j|^2)}{w_{jj}^2}R_{j,2}^2, \end{aligned}$$

where the inequality is derived using Bunjakowski-Schwarz. □

The theorem tells us that $R_{j,1}$ and $R_{j,2}$ may be interpreted as relevance measures of the j-th input variable with respect to the considered network output. If e.g., $R_{j,1}$ is smaller than either $|w_{jj}|$ or the mean of $\left|\frac{\partial f(y,\theta)}{\partial w_{jj}}\right|$ or if both expressions are small, then the j-th hidden neuron, describing the dependency of the network output on the variable y_j is negligible.

For the sake of completeness, we state this qualitative property which is also applied in other parts of neural network theory, as a lemma.

Lemma 2.6. *Under the conditions of Theorem 2.5, the j-th variable y_j has little influence on the network output, if the derivative $\frac{\partial f(y,\theta)}{\partial y_j}$ is small in average measured by either $R_{j,1}$ or $R_{j,2}$.*

Remark 1. *In the case of ϕ being the identity, the neural network function reduces to*

$$f(y,\theta) = v_0 + \sum_{j=1}^{d} v_j w_{0j} + \sum_{j=1}^{d} v_j w_{jj} y_j.$$

In that case, $R_{j,1}$ becomes $|v_j \cdot w_{jj}|$. So $R_{j,1}$ is the coefficient of the input variable y_j. Obviously, the influence of y_j for the output is small, whenever $R_{j,1}$ is small and vice versa. This therefore explains the term "relevance measure" for $R_{j,1}$.

2.4.2 Relevance measures and partial derivative plots

As already discussed in the previous sections relevance measures estimated from regression functions can be used to determine the impact of every single input variable with respect to the considered output.

In the linear case the partial derivatives coincide with the regression coefficients and thus are constants. In the case of a general nonlinear differentiable regression model the computation of the relevance measures also is possible, however in order to guarantee the interpretability of the results further structural properties have to be fulfilled. For AN models, like the considered blocked neural networks, these properties hold (see section 2.4) and therefore the impacts of the single input variables can be estimated.

In the following it will be explained, how a relevance measure and the partial derivative plots are derived and how they are interpreted.

In order to compute the different relevance measures, the first partial derivatives $\frac{\partial f()}{\partial y_j}$ of the trained network function $f(\cdot, \Theta)$ with respect to each explaining variable y_j, $j \in \{1, \ldots, d\}$ have to be determined. For the sigmoid neuron activation function of the network, the partial derivatives are calculated via:

$$\frac{\partial f(y_1, \ldots, y_d, \Theta)}{\partial y_j} = \sum_{i=H(j-1)+1}^{H(j)} v_i \left(1 - tanh^2(b_i + w_i y_j)\right) w_i, \quad j = 1, \ldots, d, \quad (6)$$

where $H(j)$ is defined as in equation (5). Obviously, the partial derivative explicitly only depends on the considered input itself, the influences of the other variables is comprised in the network parameters Θ.

The partial derivative (PD) themselves already can be used to analyze the impact of the input variables. Therefore, for each of the input variables a plot is generated from the corresponding partial derivative, that is evaluated for each given data pair. A large PD-value indicates that the influence of the related input variable is strong for the considered output value, already small changes of the input value will cause large changes on the output value. Vice versa a small PD-value is an indicator for a weak dependency. Moreover, if the PD-values for a certain input variable are small for all input values, then it is not considered to be an explaining variable, i.e. it is redundant. If for a certain range of the output all PD-values of all model input variables are small, then there is no clear causal relationship between the inputs and the output at this range. One reason for such an observation might be a missing input variable. Furthermore, a positive PD-value indicates that an increase in the input value will lead to an increase in the output value, whereas a negative PD-value indicates that an increase in the input will lead to a decrease in the output. Although each of the PD-values contains relevant information concerning the impact the related input variable, due to outliers in the data set those interpretations could be erroneous. Therefore one always should consider the PD-values of a complete input interval, where one especially should focus on those ranges with a sufficient number of data available.

Based on equation (6), the chosen relevance measure is computed by evaluating the corresponding equation for the given data set. In the following we especially focus on the AD and AvE relevance measure.
The AD relevance measure estimates the mean influence of the input variable. This number however should be interpreted with care, since i.e. a small AD-value could be the result of the sum of large positive and large negative PD-values. This measure also describes how changes in the output and the input values are correlated in the mean. A positive AD-value indicates that in the mean a positive change of the input value will increase the output value whereas a negative AD-value indicates that the output value will decrease under such input changes. The average elasticity (AvE) determines the average percentage change in the output value assuming a one percent change in each of the input values.

3 Quantification and analysis of an input output relationship in the software development process

In this section we apply the presented methods to an inspection process (coding phase), being a part of the overall software development.

Figure 1 shows the qualitative models related to this process, i.e. the control diagram, the flow diagram and the cause-effect diagram.

On the way to an implementation of the inspection process in terms of a discrete event simulation model, one important step is the quantification of the occurring input output relationships given as nodes in the related cause-effect diagram (see figure 1). In the following we focus on one of those nodes with *Number of detected major defects in an inspection* being the explained variable and *effort*, *size of the product (LOC)* and *inspected size of the product* being the explaining variables. Further input variables, that according to the cause-effect diagram also are influencing the chosen output, cannot be considered since no measurement data is available.

A linear regression model and a blocked neural network with one neuron in each block in the hidden layer were both trained based on the available data set. The cross validation performance (leave one out) of both models is shown in figure 5 and figure 6. Especially by considering the error plots for each model (see figure 6), one notices that the performance of the blocked neural network is much better compared to the linear regression model. This observation is confirmed by comparing the mean absolute errors, which is 0.8 for the blocked neural network and 1.2 for the linear regression model. This means that the neural network produces in the mean a prediction error of 0.8 *major defects* per document, compared to 1.2 *major defects* for the linear regression model. Thus, the nonlinear approach should be used to quantify the input output relationship at the considered node. Based on the existing qualitative knowledge one would expect that the performance of both models could be increased if the skipped

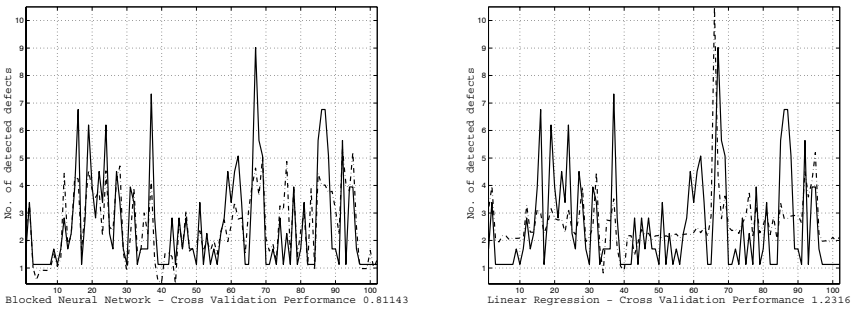

Figure 5: Left: Performance of the blocked neural network. Right: Performance of the linear regression model. Dashed line: Prediction. Solid line: Measurement

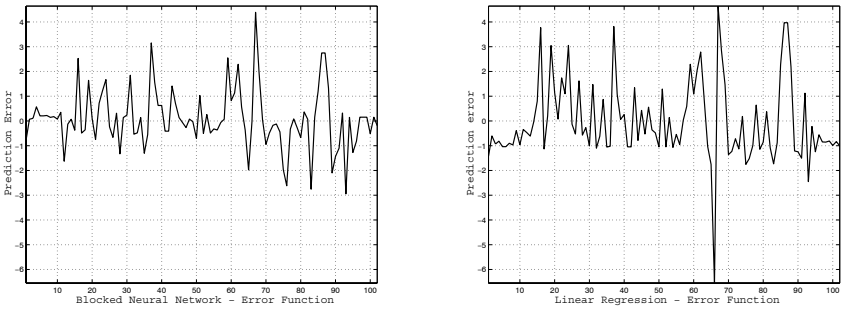

Figure 6: Left: Error between network prediction and measurement. Right: Error between linear prediction and measurement.

variables like *human effects*, *complexity* or *familiarity with the product* were also used as input variables of the models. The trained network function now can be plugged into the simulation model and then can be used to determine the *number of detected major defects found in an inspection* during the simulation runs.

Based on the trained blocked neural network we now compute and interpret the partial derivatives and the relevance measures AD and AvE. The stability of the calculated PDs for the considered neural network was proven by retraining the neural network several times.

Figure 7 shows the plots of the partial derivatives of the variable *major defects* with respect to all used input variables . One observes that the variables *effort* and *inspected size of the product* both have positive PD-values over their

whole range, while the variable *size of the product* only possesses negative PD-values.

Considering the plot for the variable *effort* in more detail one notices that for an actual effort in the range of 750 to 850 units increasing the effort while leaving the remaining input variable unchanged leads to a significant increase in the number of found defects. Obviously, the largest benefit for an increase in the working effort in terms of additionally found major defects is obtained around 775 units. An increase of the effort for documents with an actual value already greater than 850 units only will lead to a slight increase in the number of found defects, i.e. a saturation effect occurs. Thus, based on the PD-plot for the variable effort and the known costs for each effort unit a software manager approximately can determine the effort he would like to spend for the inspection. An analogous behavior of the partial derivatives can be observed for the variable *inspected size of the product*. In contrast to the already considered two input variables the explaining variable *size of the product* has negative or zero partial derivatives. This means leaving the variable *effort* and *inspected size of the product* unchanged and increasing the *size of the product* leads to a smaller number of detected defects. One has to keep in mind that in this case smaller percentage of the document will be inspected and that the effort and the inspected lines of code are unchanged.

Figure 8 depicts of the AD-relevance measures (2) of the inputs, whereas figure 9 displays the corresponding AvE-measures (see equation (3)). In both figures one observes that the variable *effort* has the largest impact with respect to the variable *detected major defects in the inspection*. As explained in the last section the mean sign of correlation between the inputs and the output can be determined by the AD-value, which is positive for the variables *effort* and *inspected size of the product* and negative for the variable *size of the product*.

All in all, the partial derivatives as well as the relevance measures contain important quantitative information about the influence of the input variables with respect to the considered output and thus provide the software manager with a more detailed insight into the structure of the input output relation. He especially gets able to estimate the impact of changes in the process.

In general the partial derivatives, the AD and AvE relevance measures also can be used to validate a cause-effect diagram. Due to the lack of data this aspect is not considered here in detail. In general the derived methodology allows to check whether a variable is missing or redundant in the cause-effect diagram, or if the signs indicating the direction of correlation between the input and output are correct for the considered output node.

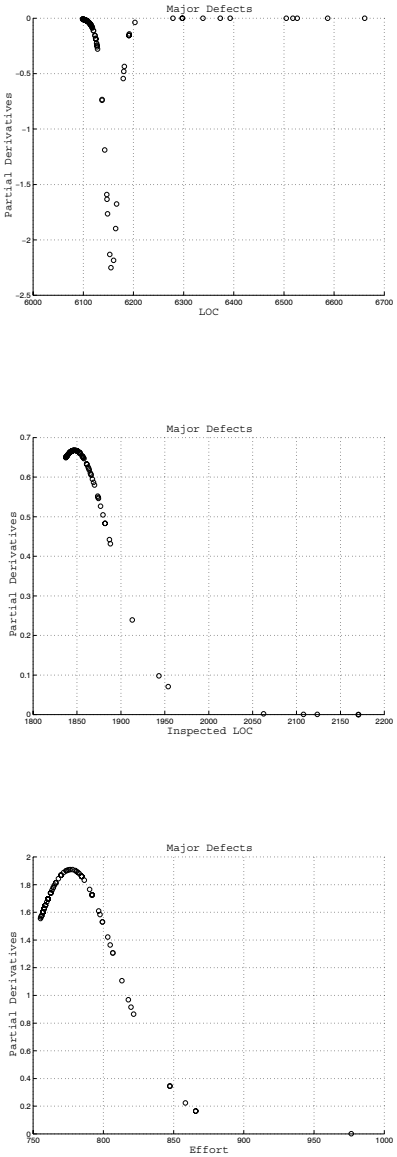

Figure 7: Plot of partial derivatives for the variable *Number of detected defects in an inspection* with respect to the variables *size of the product* (top), *inspected size of the product* (middle), *effort* (bottom)

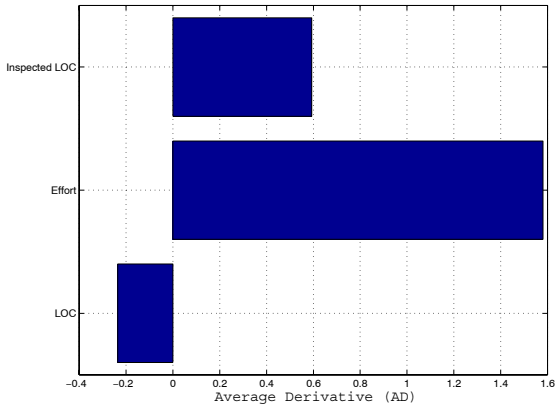

Figure 8: Plot of Average Derivative (AD)

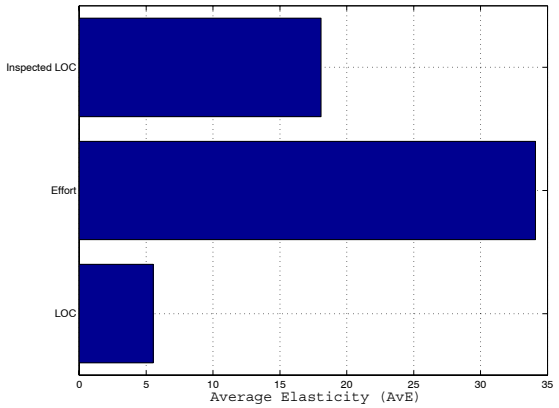

Figure 9: Plot of Average Elasticity (AE)

4 Conclusion

In this paper we have presented blocked neural networks as an appropriate tool for quantification of static input output relationships of software development processes. The performance of a blocked neural network with one neuron in each block turned out to be superior compared to a linear regression model. The major advantage of blocked networks compared to other nonlinear regression models is the interpretability of the related partial derivatives and relevance measures with respect to the impact of the single input variables. Plots of the partial derivatives can be used to estimate the functional influence of each input variable with respect to the output and thus can be used to find new rules of thumb. They also can be used for validation of the existing qualitative models of a software development process.

5 Acknowledgements

The authors gratefully acknowledge the support of this research done in connection with the SEV project supported by the German Bundesministerium für Bildung und Forschung and the Prosim project supported by the Stiftung Rheinlandpfalz für Innovation, project no. 559.

References

[1] S. M. Carroll, B. W. Dickinson: *Construction of neural nets using the Radon transform.* '1989 IJCNN Proceedings.

[2] G. Cybenko: *Approximation by superpositions of sigmoidal functions.* Math. Control, Signals, Systems, 2, 1989.

[3] B. Efron, R. J Tibshirani: *An Introduction to the Bootstrap*, Monographs on Statistics and Applied Probability 57, Chapman & Hall Inc., New York, 1993

[4] Ken-ichi Funahashi: *On the Approximate Realization of Continuous Mappings by Neural Networks.* Neural Networks, 2(1989).

[5] Meihui Guo, Zhidong Bai, Hong Zhi: *Multi-Step prediction for Nonlinear Autoregressive Models Based On Empirical Distributions.* Statistica Sinica 9(1999).

[6] J. Münch, H. Neu, T. Berlage, T. Hanne, S. Nickel; S. Von Stockum; A. Wirsen: *Simulation-based Evaluation and Improvement of Software Developement Processes*,IESE-Report 048.02/E, August 2002

[7] H. Neu, T. Hanne, J. Münch, S. Nickel, A. Wirsen: *Simulation Based Risk Reduction for Planning Inspections*, Berlin, Springer Verlag, 2002,xv,646 P, Lecture Notes in Computer Science, 2559, Inter-

national Conference on Product Focused Software Improvement, 2002 Rovaniemi.

[8] T. Hastie, R. Tibshirani, J. Friedman: *The Elements of Statistical Learning.*, Springer 2001

[9] K. Hornik, M. Stinchcombe, H. White: *Multileyer feedforward networks are universal approximators.* Neural Networks 2(1989).

[10] K. Hornik, M. Stinchcombe, H. White: *Universal Approximation of an Unknown Mapping and Its Derivatives Using Multilayer Feedforward Networks.* Neural Networks 3(1990).

[11] W. Y. Huang, R. Lippmann: *Neural Net and Traditional Classifiers*, NIPS 1987.

[12] M. Lehtokangas, J. Saarinen, K. Kaski, P. Huuhtanen: *A Network of Autoregressive Processing Units for Time Series Modeling.* Applied Mathematics and Computation, 75, 1996.

[13] R. P. Lippmann: *An introduction to computing with neural nets.* ILEE 1987.

[14] B. M. Pötscher, I. R. Prucha: *Dynamic Nonlinear Econometric Models, Asymptotic Theory.* Springer 1997.

[15] P. N. Refenes, A. D. Zapranis, J. Utans: *Neural model identification, variable selection and model adequacy.* Neural Networks in Financial Engineering, Proc. NnCM-1996.

[16] A. Sarishvili: *Neural Network Based Lack Selection, for multivariate time series*, Dissertation Universit"at Kaiserslautern, 2002

[17] Tschernig R, Yang LJ. : *Nonparametric lag selection for time series.* Journal of Time Series Analysis 21:(4) Jul 2000.

[18] H. White: *Learning in artificial neural networks: A statistical perspective.* Neural Computation 1(1989).

[19] Qiwei Yao, Howell Tong: *On subset selection in non-parametric stochastic regression* Statistica Sinca 4(1994).

[20] Ning Xiong: *Designing Compact and Comprehensible Fuzzy Controllers Using Genetic Algorithms*, Dissertation University of Kaiserslautern, Shaker Verlag, Aachen 2001.

Published reports of the Fraunhofer ITWM

The PDF-files of the following reports are available under:
www.itwm.fraunhofer.de/rd/presse/berichte

1. D. Hietel, K. Steiner, J. Struckmeier

A Finite - Volume Particle Method for Compressible Flows

We derive a new class of particle methods for conservation laws, which are based on numerical flux functions to model the interactions between moving particles. The derivation is similar to that of classical Finite-Volume methods; except that the fixed grid structure in the Finite-Volume method is substituted by so-called mass packets of particles. We give some numerical results on a shock wave solution for Burgers equation as well as the well-known one-dimensional shock tube problem.
(19 pages, 1998)

2. M. Feldmann, S. Seibold

Damage Diagnosis of Rotors: Application of Hilbert Transform and Multi-Hypothesis Testing

In this paper, a combined approach to damage diagnosis of rotors is proposed. The intention is to employ signal-based as well as model-based procedures for an improved detection of size and location of the damage. In a first step, Hilbert transform signal processing techniques allow for a computation of the signal envelope and the instantaneous frequency, so that various types of non-linearities due to a damage may be identified and classified based on measured response data. In a second step, a multi-hypothesis bank of Kalman Filters is employed for the detection of the size and location of the damage based on the information of the type of damage provided by the results of the Hilbert transform.
Keywords: Hilbert transform, damage diagnosis, Kalman filtering, non-linear dynamics
(23 pages, 1998)

3. Y. Ben-Haim, S. Seibold

Robust Reliability of Diagnostic Multi-Hypothesis Algorithms: Application to Rotating Machinery

Damage diagnosis based on a bank of Kalman filters, each one conditioned on a specific hypothesized system condition, is a well recognized and powerful diagnostic tool. This multi-hypothesis approach can be applied to a wide range of damage conditions. In this paper, we will focus on the diagnosis of cracks in rotating machinery. The question we address is: how to optimize the multi-hypothesis algorithm with respect to the uncertainty of the spatial form and location of cracks and their resulting dynamic effects. First, we formulate a measure of the reliability of the diagnostic algorithm, and then we discuss modifications of the diagnostic algorithm for the maximization of the reliability. The reliability of a diagnostic algorithm is measured by the amount of uncertainty consistent with no-failure of the diagnosis. Uncertainty is quantitatively represented with convex models.
Keywords: Robust reliability, convex models, Kalman filtering, multi-hypothesis diagnosis, rotating machinery, crack diagnosis
(24 pages, 1998)

4. F.-Th. Lentes, N. Siedow

Three-dimensional Radiative Heat Transfer in Glass Cooling Processes

For the numerical simulation of 3D radiative heat transfer in glasses and glass melts, practically applicable mathematical methods are needed to handle such problems optimal using workstation class computers. Since the exact solution would require super-computer capabilities we concentrate on approximate solutions with a high degree of accuracy. The following approaches are studied: 3D diffusion approximations and 3D ray-tracing methods.
(23 pages, 1998)

5. A. Klar, R. Wegener

A hierarchy of models for multilane vehicular traffic
Part I: Modeling

In the present paper multilane models for vehicular traffic are considered. A microscopic multilane model based on reaction thresholds is developed. Based on this model an Enskog like kinetic model is developed. In particular, care is taken to incorporate the correlations between the vehicles. From the kinetic model a fluid dynamic model is derived. The macroscopic coefficients are deduced from the underlying kinetic model. Numerical simulations are presented for all three levels of description in [10]. Moreover, a comparison of the results is given there.
(23 pages, 1998)

Part II: Numerical and stochastic investigations

In this paper the work presented in [6] is continued. The present paper contains detailed numerical investigations of the models developed there. A numerical method to treat the kinetic equations obtained in [6] are presented and results of the simulations are shown. Moreover, the stochastic correlation model used in [6] is described and investigated in more detail.
(17 pages, 1998)

6. A. Klar, N. Siedow

Boundary Layers and Domain Decomposition for Radiative Heat Transfer and Diffusion Equations: Applications to Glass Manufacturing Processes

In this paper domain decomposition methods for radiative transfer problems including conductive heat transfer are treated. The paper focuses on semi-transparent materials, like glass, and the associated conditions at the interface between the materials. Using asymptotic analysis we derive conditions for the coupling of the radiative transfer equations and a diffusion approximation. Several test cases are treated and a problem appearing in glass manufacturing processes is computed. The results clearly show the advantages of a domain decomposition approach. Accuracy equivalent to the solution of the global radiative transfer solution is achieved, whereas computation time is strongly reduced.
(24 pages, 1998)

7. I. Choquet

Heterogeneous catalysis modelling and numerical simulation in rarified gas flows Part I: Coverage locally at equilibrium

A new approach is proposed to model and simulate numerically heterogeneous catalysis in rarefied gas flows. It is developed to satisfy all together the following points:
1) describe the gas phase at the microscopic scale, as required in rarefied flows,
2) describe the wall at the macroscopic scale, to avoid prohibitive computational costs and consider not only crystalline but also amorphous surfaces,
3) reproduce on average macroscopic laws correlated with experimental results and
4) derive analytic models in a systematic and exact way. The problem is stated in the general framework of a non static flow in the vicinity of a catalytic and non porous surface (without aging). It is shown that the exact and systematic resolution method based on the Laplace transform, introduced previously by the author to model collisions in the gas phase, can be extended to the present problem. The proposed approach is applied to the modelling of the EleyRideal and LangmuirHinshelwood recombinations, assuming that the coverage is locally at equilibrium. The models are developed considering one atomic species and extended to the general case of several atomic species. Numerical calculations show that the models derived in this way reproduce with accuracy behaviors observed experimentally.
(24 pages, 1998)

8. J. Ohser, B. Steinbach, C. Lang

Efficient Texture Analysis of Binary Images

A new method of determining some characteristics of binary images is proposed based on a special linear filtering. This technique enables the estimation of the area fraction, the specific line length, and the specific integral of curvature. Furthermore, the specific length of the total projection is obtained, which gives detailed information about the texture of the image. The influence of lateral and directional resolution depending on the size of the applied filter mask is discussed in detail. The technique includes a method of increasing directional resolution for texture analysis while keeping lateral resolution as high as possible.
(17 pages, 1998)

9. J. Orlik

Homogenization for viscoelasticity of the integral type with aging and shrinkage

A multiphase composite with periodic distributed inclusions with a smooth boundary is considered in this contribution. The composite component materials are supposed to be linear viscoelastic and aging (of the nonconvolution integral type, for which the Laplace transform with respect to time is not effectively applicable) and are subjected to isotropic shrinkage. The free shrinkage deformation can be considered as a fictitious temperature deformation in the behavior law. The procedure presented in this paper proposes a way to determine average (effective homogenized) viscoelastic and shrinkage (temperature) composite properties and the homogenized stressfield from known properties of the components. This is done by the extension of the asymptotic homogenization technique known for pure elastic nonhomogeneous bodies to the nonhomogeneous thermoviscoelasticity of the integral noncon-

volution type. Up to now, the homogenization theory has not covered viscoelasticity of the integral type. SanchezPalencia (1980), Francfort & Suquet (1987) (see [2], [9]) have considered homogenization for viscoelasticity of the differential form and only up to the first derivative order. The integralmodeled viscoelasticity is more general then the differential one and includes almost all known differential models. The homogenization procedure is based on the construction of an asymptotic solution with respect to a period of the composite structure. This reduces the original problem to some auxiliary boundary value problems of elasticity and viscoelasticity on the unit periodic cell, of the same type as the original non-homogeneous problem. The existence and uniqueness results for such problems were obtained for kernels satisfying some constrain conditions. This is done by the extension of the Volterra integral operator theory to the Volterra operators with respect to the time, whose 1 kernels are space linear operators for any fixed time variables. Some ideas of such approach were proposed in [11] and [12], where the Volterra operators with kernels depending additionally on parameter were considered. This manuscript delivers results of the same nature for the case of the spaceoperator kernels.
(20 pages, 1998)

10. J. Mohring

Helmholtz Resonators with Large Aperture

The lowest resonant frequency of a cavity resonator is usually approximated by the classical Helmholtz formula. However, if the opening is rather large and the front wall is narrow this formula is no longer valid. Here we present a correction which is of third order in the ratio of the diameters of aperture and cavity. In addition to the high accuracy it allows to estimate the damping due to radiation. The result is found by applying the method of matched asymptotic expansions. The correction contains form factors describing the shapes of opening and cavity. They are computed for a number of standard geometries. Results are compared with numerical computations.
(21 pages, 1998)

11. H. W. Hamacher, A. Schöbel

On Center Cycles in Grid Graphs

Finding "good" cycles in graphs is a problem of great interest in graph theory as well as in locational analysis. We show that the center and median problems are NP hard in general graphs. This result holds both for the variable cardinality case (i.e. all cycles of the graph are considered) and the fixed cardinality case (i.e. only cycles with a given cardinality p are feasible). Hence it is of interest to investigate special cases where the problem is solvable in polynomial time. In grid graphs, the variable cardinality case is, for instance, trivially solvable if the shape of the cycle can be chosen freely. If the shape is fixed to be a rectangle one can analyze rectangles in grid graphs with, in sequence, fixed dimension, fixed cardinality, and variable cardinality. In all cases a complete characterization of the optimal cycles and closed form expressions of the optimal objective values are given, yielding polynomial time algorithms for all cases of center rectangle problems. Finally, it is shown that center cycles can be chosen as rectangles for small cardinalities such that the center cycle problem in grid graphs is in these cases completely solved.
(15 pages, 1998)

12. H. W. Hamacher, K.-H. Küfer

Inverse radiation therapy planning - a multiple objective optimisation approach

For some decades radiation therapy has been proved successful in cancer treatment. It is the major task of clinical radiation treatment planning to realize on the one hand a high level dose of radiation in the cancer tissue in order to obtain maximum tumor control. On the other hand it is obvious that it is absolutely necessary to keep in the tissue outside the tumor, particularly in organs at risk, the unavoidable radiation as low as possible.
No doubt, these two objectives of treatment planning - high level dose in the tumor, low radiation outside the tumor - have a basically contradictory nature. Therefore, it is no surprise that inverse mathematical models with dose distribution bounds tend to be infeasible in most cases. Thus, there is need for approximations compromising between overdosing the organs at risk and underdosing the target volume.
Differing from the currently used time consuming iterative approach, which measures deviation from an ideal (non-achievable) treatment plan using recursively trial-and-error weights for the organs of interest, we go a new way trying to avoid a priori weight choices and consider the treatment planning problem as a multiple objective linear programming problem: with each organ of interest, target tissue as well as organs at risk, we associate an objective function measuring the maximal deviation from the prescribed doses.
We build up a data base of relatively few efficient solutions representing and approximating the variety of Pareto solutions of the multiple objective linear programming problem. This data base can be easily scanned by physicians looking for an adequate treatment plan with the aid of an appropriate online tool.
(14 pages, 1999)

13. C. Lang, J. Ohser, R. Hilfer

On the Analysis of Spatial Binary Images

This paper deals with the characterization of microscopically heterogeneous, but macroscopically homogeneous spatial structures. A new method is presented which is strictly based on integral-geometric formulae such as Crofton's intersection formulae and Hadwiger's recursive definition of the Euler number. The corresponding algorithms have clear advantages over other techniques. As an example of application we consider the analysis of spatial digital images produced by means of Computer Assisted Tomography.
(20 pages, 1999)

14. M. Junk

On the Construction of Discrete Equilibrium Distributions for Kinetic Schemes

A general approach to the construction of discrete equilibrium distributions is presented. Such distribution functions can be used to set up Kinetic Schemes as well as Lattice Boltzmann methods. The general principles are also applied to the construction of Chapman Enskog distributions which are used in Kinetic Schemes for compressible Navier-Stokes equations.
(24 pages, 1999)

15. M. Junk, S. V. Raghurame Rao

A new discrete velocity method for Navier-Stokes equations

The relation between the Lattice Boltzmann Method, which has recently become popular, and the Kinetic Schemes, which are routinely used in Computational Fluid Dynamics, is explored. A new discrete velocity model for the numerical solution of Navier-Stokes equations for incompressible fluid flow is presented by combining both the approaches. The new scheme can be interpreted as a pseudo-compressibility method and, for a particular choice of parameters, this interpretation carries over to the Lattice Boltzmann Method.
(20 pages, 1999)

16. H. Neunzert

Mathematics as a Key to Key Technologies

The main part of this paper will consist of examples, how mathematics really helps to solve industrial problems; these examples are taken from our Institute for Industrial Mathematics, from research in the Technomathematics group at my university, but also from ECMI groups and a company called TecMath, which originated 10 years ago from my university group and has already a very successful history.
(39 pages (4 PDF-Files), 1999)

17. J. Ohser, K. Sandau

Considerations about the Estimation of the Size Distribution in Wicksell's Corpuscle Problem

Wicksell's corpuscle problem deals with the estimation of the size distribution of a population of particles, all having the same shape, using a lower dimensional sampling probe. This problem was originary formulated for particle systems occurring in life sciences but its solution is of actual and increasing interest in materials science. From a mathematical point of view, Wicksell's problem is an inverse problem where the interesting size distribution is the unknown part of a Volterra equation. The problem is often regarded ill-posed, because the structure of the integrand implies unstable numerical solutions. The accuracy of the numerical solutions is considered here using the condition number, which allows to compare different numerical methods with different (equidistant) class sizes and which indicates, as one result, that a finite section thickness of the probe reduces the numerical problems. Furthermore, the relative error of estimation is computed which can be split into two parts. One part consists of the relative discretization error that increases for increasing class size, and the second part is related to the relative statistical error which increases with decreasing class size. For both parts, upper bounds can be given and the sum of them indicates an optimal class width depending on some specific constants.
(18 pages, 1999)

18. E. Carrizosa, H. W. Hamacher, R. Klein, S. Nickel

Solving nonconvex planar location problems by finite dominating sets

It is well-known that some of the classical location problems with polyhedral gauges can be solved in polynomial time by finding a finite dominating set, i.e. a finite set of candidates guaranteed to contain at least one optimal location.
In this paper it is first established that this result holds

for a much larger class of problems than currently considered in the literature. The model for which this result can be proven includes, for instance, location problems with attraction and repulsion, and location-allocation problems.
Next, it is shown that the approximation of general gauges by polyhedral ones in the objective function of our general model can be analyzed with regard to the subsequent error in the optimal objective value. For the approximation problem two different approaches are described, the sandwich procedure and the greedy algorithm. Both of these approaches lead - for fixed epsilon - to polynomial approximation algorithms with accuracy epsilon for solving the general model considered in this paper.
Keywords: Continuous Location, Polyhedral Gauges, Finite Dominating Sets, Approximation, Sandwich Algorithm, Greedy Algorithm
(19 pages, 2000)

19. A. Becker

A Review on Image Distortion Measures

Within this paper we review image distortion measures. A distortion measure is a criterion that assigns a "quality number" to an image. We distinguish between mathematical distortion measures and those distortion measures in-cooperating a priori knowledge about the imaging devices (e. g. satellite images), image processing algorithms or the human physiology. We will consider representative examples of different kinds of distortion measures and are going to discuss them.
Keywords: Distortion measure, human visual system
(26 pages, 2000)

20. H. W. Hamacher, M. Labbé, S. Nickel, T. Sonneborn

Polyhedral Properties of the Uncapacitated Multiple Allocation Hub Location Problem

We examine the feasibility polyhedron of the uncapacitated hub location problem (UHL) with multiple allocation, which has applications in the fields of air passenger and cargo transportation, telecommunication and postal delivery services. In particular we determine the dimension and derive some classes of facets of this polyhedron. We develop some general rules about lifting facets from the uncapacitated facility location (UFL) for UHL and projecting facets from UHL to UFL. By applying these rules we get a new class of facets for UHL which dominates the inequalities in the original formulation. Thus we get a new formulation of UHL whose constraints are all facet–defining. We show its superior computational performance by benchmarking it on a well known data set.
Keywords: integer programming, hub location, facility location, valid inequalities, facets, branch and cut
(21 pages, 2000)

21. H. W. Hamacher, A. Schöbel

Design of Zone Tariff Systems in Public Transportation

Given a public transportation system represented by its stops and direct connections between stops, we consider two problems dealing with the prices for the customers: The fare problem in which subsets of stops are already aggregated to zones and "good" tariffs have to be found in the existing zone system. Closed form solutions for the fare problem are presented for three objective functions. In the zone problem the design of the zones is part of the problem. This problem is NP hard and we therefore propose three heuristics which prove to be very successful in the redesign of one of Germany's transportation systems.
(30 pages, 2001)

22. D. Hietel, M. Junk, R. Keck, D. Teleaga:

The Finite-Volume-Particle Method for Conservation Laws

In the Finite-Volume-Particle Method (FVPM), the weak formulation of a hyperbolic conservation law is discretized by restricting it to a discrete set of test functions. In contrast to the usual Finite-Volume approach, the test functions are not taken as characteristic functions of the control volumes in a spatial grid, but are chosen from a partition of unity with smooth and overlapping partition functions (the particles), which can even move along pre- scribed velocity fields. The information exchange between particles is based on standard numerical flux functions. Geometrical information, similar to the surface area of the cell faces in the Finite-Volume Method and the corresponding normal directions are given as integral quantities of the partition functions. After a brief derivation of the Finite-Volume-Particle Method, this work focuses on the role of the geometric coefficients in the scheme.
(16 pages, 2001)

23. T. Bender, H. Hennes, J. Kalcsics, M. T. Melo, S. Nickel

Location Software and Interface with GIS and Supply Chain Management

The objective of this paper is to bridge the gap between location theory and practice. To meet this objective focus is given to the development of software capable of addressing the different needs of a wide group of users. There is a very active community on location theory encompassing many research fields such as operations research, computer science, mathematics, engineering, geography, economics and marketing. As a result, people working on facility location problems have a very diverse background and also different needs regarding the software to solve these problems. For those interested in non-commercial applications (e. g. students and researchers), the library of location algorithms (LoLA can be of considerable assistance. LoLA contains a collection of efficient algorithms for solving planar, network and discrete facility location problems. In this paper, a detailed description of the functionality of LoLA is presented. In the fields of geography and marketing, for instance, solving facility location problems requires using large amounts of demographic data. Hence, members of these groups (e. g. urban planners and sales managers) often work with geographical information too s. To address the specific needs of these users, LoLA was inked to a geographical information system (GIS) and the details of the combined functionality are described in the paper. Finally, there is a wide group of practitioners who need to solve large problems and require special purpose software with a good data interface. Many of such users can be found, for example, in the area of supply chain management (SCM). Logistics activities involved in strategic SCM include, among others, facility location planning. In this paper, the development of a commercial location software tool is also described. The too is embedded in the Advanced Planner and Optimizer SCM software developed by SAP AG, Walldorf, Germany. The paper ends with some conclusions and an outlook to future activities.
Keywords: facility location, software development, geographical information systems, supply chain management.
(48 pages, 2001)

24. H. W. Hamacher, S. A. Tjandra

Mathematical Modelling of Evacuation Problems: A State of Art

This paper details models and algorithms which can be applied to evacuation problems. While it concentrates on building evacuation many of the results are applicable also to regional evacuation. All models consider the time as main parameter, where the travel time between components of the building is part of the input and the overall evacuation time is the output. The paper distinguishes between macroscopic and microscopic evacuation models both of which are able to capture the evacuees' movement over time.
Macroscopic models are mainly used to produce good lower bounds for the evacuation time and do not consider any individual behavior during the emergency situation. These bounds can be used to analyze existing buildings or help in the design phase of planning a building. Macroscopic approaches which are based on dynamic network flow models (minimum cost dynamic flow, maximum dynamic flow, universal maximum flow, quickest path and quickest flow) are described. A special feature of the presented approach is the fact, that travel times of evacuees are not restricted to be constant, but may be density dependent. Using multicriteria optimization priority regions and blockage due to fire or smoke may be considered. It is shown how the modelling can be done using time parameter either as discrete or continuous parameter.
Microscopic models are able to model the individual evacuee's characteristics and the interaction among evacuees which influence their movement. Due to the corresponding huge amount of data one uses simulation approaches. Some probabilistic laws for individual evacuee's movement are presented. Moreover ideas to model the evacuee's movement using cellular automata (CA) and resulting software are presented. In this paper we will focus on macroscopic models and only summarize some of the results of the microscopic approach. While most of the results are applicable to general evacuation situations, we concentrate on building evacuation.
(44 pages, 2001)

25. J. Kuhnert, S. Tiwari

Grid free method for solving the Poisson equation

A Grid free method for solving the Poisson equation is presented. This is an iterative method. The method is based on the weighted least squares approximation in which the Poisson equation is enforced to be satisfied in every iterations. The boundary conditions can also be enforced in the iteration process. This is a local approximation procedure. The Dirichlet, Neumann and mixed boundary value problems on a unit square are presented and the analytical solutions are compared with the exact solutions. Both solutions matched perfectly.
Keywords: Poisson equation, Least squares method, Grid free method
(19 pages, 2001)

26. T. Götz, H. Rave, D. Reinel-Bitzer, K. Steiner, H. Tiemeier

Simulation of the fiber spinning process

To simulate the influence of process parameters to the melt spinning process a fiber model is used and coupled with CFD calculations of the quench air flow. In the fiber model energy, momentum and mass balance are solved for the polymer mass flow. To calculate the quench air the Lattice Boltzmann method is used. Simulations and experiments for different process parameters and hole configurations are compared and show a good agreement.

Keywords: Melt spinning, fiber model, Lattice Boltzmann, CFD

(19 pages, 2001)

27. A. Zemitis

On interaction of a liquid film with an obstacle

In this paper mathematical models for liquid films generated by impinging jets are discussed. Attention is stressed to the interaction of the liquid film with some obstacle. S. G. Taylor [Proc. R. Soc. London Ser. A 253, 313 (1959)] found that the liquid film generated by impinging jets is very sensitive to properties of the wire which was used as an obstacle. The aim of this presentation is to propose a modification of the Taylor's model, which allows to simulate the film shape in cases, when the angle between jets is different from 180°. Numerical results obtained by discussed models give two different shapes of the liquid film similar as in Taylors experiments. These two shapes depend on the regime: either droplets are produced close to the obstacle or not. The difference between two regimes becomes larger if the angle between jets decreases. Existence of such two regimes can be very essential for some applications of impinging jets, if the generated liquid film can have a contact with obstacles.

Keywords: impinging jets, liquid film, models, numerical solution, shape

(22 pages, 2001)

28. I. Ginzburg, K. Steiner

Free surface lattice-Boltzmann method to model the filling of expanding cavities by Bingham Fluids

The filling process of viscoplastic metal alloys and plastics in expanding cavities is modelled using the lattice Boltzmann method in two and three dimensions. These models combine the regularized Bingham model for viscoplastic with a free-interface algorithm. The latter is based on a modified immiscible lattice Boltzmann model in which one species is the fluid and the other one is considered as vacuum. The boundary conditions at the curved liquid-vacuum interface are met without any geometrical front reconstruction from a first-order Chapman-Enskog expansion. The numerical results obtained with these models are found in good agreement with available theoretical and numerical analysis.

Keywords: Generalized LBE, free-surface phenomena, interface boundary conditions, filling processes, Bingham viscoplastic model, regularized models

(22 pages, 2001)

29. H. Neunzert

»Denn nichts ist für den Menschen als Menschen etwas wert, was er nicht mit Leidenschaft tun kann«

Vortrag anlässlich der Verleihung des Akademiepreises des Landes Rheinland-Pfalz am 21.11.2001

Was macht einen guten Hochschullehrer aus? Auf diese Frage gibt es sicher viele verschiedene, fachbezogene Antworten, aber auch ein paar allgemeine Gesichtspunkte: es bedarf der »Leidenschaft« für die Forschung (Max Weber), aus der dann auch die Begeisterung für die Lehre erwächst. Forschung und Lehre gehören zusammen, um die Wissenschaft als lebendiges Tun vermitteln zu können. Der Vortrag gibt Beispiele dafür, wie in angewandter Mathematik Forschungsaufgaben aus praktischen Alltagsproblemstellungen erwachsen, die in die Lehre auf verschiedenen Stufen (Gymnasium bis Graduiertenkolleg) einfließen; er leitet damit auch zu einem aktuellen Forschungsgebiet, der Mehrskalenanalyse mit ihren vielfältigen Anwendungen in Bildverarbeitung, Materialentwicklung und Strömungsmechanik über, was aber nur kurz gestreift wird. Mathematik erscheint hier als eine moderne Schlüsseltechnologie, die aber auch enge Beziehungen zu den Geistes- und Sozialwissenschaften hat.

Keywords: Lehre, Forschung, angewandte Mathematik, Mehrskalenanalyse, Strömungsmechanik

(18 pages, 2001)

30. J. Kuhnert, S. Tiwari

Finite pointset method based on the projection method for simulations of the incompressible Navier-Stokes equations

A Lagrangian particle scheme is applied to the projection method for the incompressible Navier-Stokes equations. The approximation of spatial derivatives is obtained by the weighted least squares method. The pressure Poisson equation is solved by a local iterative procedure with the help of the least squares method. Numerical tests are performed for two dimensional cases. The Couette flow, Poiseuelle flow, decaying shear flow and the driven cavity flow are presented. The numerical solutions are obtained for stationary as well as instationary cases and are compared with the analytical solutions for channel flows. Finally, the driven cavity in a unit square is considered and the stationary solution obtained from this scheme is compared with that from the finite element method.

Keywords: Incompressible Navier-Stokes equations, Meshfree method, Projection method, Particle scheme, Least squares approximation

AMS subject classification: 76D05, 76M28

(25 pages, 2001)

31. R. Korn, M. Krekel

Optimal Portfolios with Fixed Consumption or Income Streams

We consider some portfolio optimisation problems where either the investor has a desire for an a priori specified consumption stream or/and follows a deterministic pay in scheme while also trying to maximize expected utility from final wealth. We derive explicit closed form solutions for continuous and discrete monetary streams. The mathematical method used is classical stochastic control theory.

Keywords: Portfolio optimisation, stochastic control, HJB equation, discretisation of control problems.

(23 pages, 2002)

32. M. Krekel

Optimal portfolios with a loan dependent credit spread

If an investor borrows money he generally has to pay higher interest rates than he would have received, if he had put his funds on a savings account. The classical model of continuous time portfolio optimisation ignores this effect. Since there is obviously a connection between the default probability and the total percentage of wealth, which the investor is in debt, we study portfolio optimisation with a control dependent interest rate. Assuming a logarithmic and a power utility function, respectively, we prove explicit formulae of the optimal control.

Keywords: Portfolio optimisation, stochastic control, HJB equation, credit spread, log utility, power utility, non-linear wealth dynamics

(25 pages, 2002)

33. J. Ohser, W. Nagel, K. Schladitz

The Euler number of discretized sets - on the choice of adjacency in homogeneous lattices

Two approaches for determining the Euler-Poincaré characteristic of a set observed on lattice points are considered in the context of image analysis { the integral geometric and the polyhedral approach. Information about the set is assumed to be available on lattice points only. In order to retain properties of the Euler number and to provide a good approximation of the true Euler number of the original set in the Euclidean space, the appropriate choice of adjacency in the lattice for the set and its background is crucial. Adjacencies are defined using tessellations of the whole space into polyhedrons. In R 3 , two new 14 adjacencies are introduced additionally to the well known 6 and 26 adjacencies. For the Euler number of a set and its complement, a consistency relation holds. Each of the pairs of adjacencies (14:1; 14:1), (14:2; 14:2), (6; 26), and (26; 6) is shown to be a pair of complementary adjacencies with respect to this relation. That is, the approximations of the Euler numbers are consistent if the set and its background (complement) are equipped with this pair of adjacencies. Furthermore, sufficient conditions for the correctness of the approximations of the Euler number are given. The analysis of selected microstructures and a simulation study illustrate how the estimated Euler number depends on the chosen adjacency. It also shows that there is not a uniquely best pair of adjacencies with respect to the estimation of the Euler number of a set in Euclidean space.

Keywords: image analysis, Euler number, neighborhod relationships, cuboidal lattice

(32 pages, 2002)

34. I. Ginzburg, K. Steiner

Lattice Boltzmann Model for Free-Surface flow and Its Application to Filling Process in Casting

A generalized lattice Boltzmann model to simulate free-surface is constructed in both two and three dimensions. The proposed model satisfies the interfacial boundary conditions accurately. A distinctive feature of the model is that the collision processes is carried out only on the points occupied partially or fully by the fluid. To maintain a sharp interfacial front, the method includes an anti-diffusion algorithm. The unknown distribution functions at the interfacial region are constructed according to the first order Chapman-Enskog analysis. The interfacial boundary conditions are satis-

fied exactly by the coefficients in the Chapman-Enskog expansion. The distribution functions are naturally expressed in the local interfacial coordinates. The macroscopic quantities at the interface are extracted from the least-square solutions of a locally linearized system obtained from the known distribution functions. The proposed method does not require any geometric front construction and is robust for any interfacial topology. Simulation results of realistic filling process are presented: rectangular cavity in two dimensions and Hammer box, Campbell box, Sheffield box, and Motorblock in three dimensions. To enhance the stability at high Reynolds numbers, various upwind-type schemes are developed. Free-slip and no-slip boundary conditions are also discussed.
Keywords: Lattice Boltzmann models; free-surface phenomena; interface boundary conditions; filling processes; injection molding; volume of fluid method; interface boundary conditions; advection-schemes; upwind-schemes
(54 pages, 2002)

35. M. Günther, A. Klar, T. Materne, R. Wegener

Multivalued fundamental diagrams and stop and go waves for continuum traffic equations

In the present paper a kinetic model for vehicular traffic leading to multivalued fundamental diagrams is developed and investigated in detail. For this model phase transitions can appear depending on the local density and velocity of the flow. A derivation of associated macroscopic traffic equations from the kinetic equation is given. Moreover, numerical experiments show the appearance of stop and go waves for highway traffic with a bottleneck.
Keywords: traffic flow, macroscopic equations, kinetic derivation, multivalued fundamental diagram, stop and go waves, phase transitions
(25 pages, 2002)

36. S. Feldmann, P. Lang, D. Prätzel-Wolters

Parameter influence on the zeros of network determinants

To a network N(q) with determinant D(s;q) depending on a parameter vector $q \hat{I} R^r$ via identification of some of its vertices, a network N^ (q) is assigned. The paper deals with procedures to find N^ (q), such that its determinant D^ (s;q) admits a factorization in the determinants of appropriate subnetworks, and with the estimation of the deviation of the zeros of D^ from the zeros of D. To solve the estimation problem state space methods are applied.
Keywords: Networks, Equicofactor matrix polynomials, Realization theory, Matrix perturbation theory
(30 pages, 2002)

37. K. Koch, J. Ohser, K. Schladitz

Spectral theory for random closed sets and estimating the covariance via frequency space

A spectral theory for stationary random closed sets is developed and provided with a sound mathematical basis. Definition and proof of existence of the Bartlett spectrum of a stationary random closed set as well as the proof of a Wiener-Khintchine theorem for the power spectrum are used to two ends: First, well known second order characteristics like the covariance can be estimated faster than usual via frequency space. Second, the Bartlett spectrum and the power spectrum can be used as second order characteristics in frequency space. Examples show, that in some cases information about the random closed set is easier to obtain from these characteristics in frequency space than from their real world counterparts.
Keywords: Random set, Bartlett spectrum, fast Fourier transform, power spectrum
(28 pages, 2002)

38. D. d'Humières, I. Ginzburg

Multi-reflection boundary conditions for lattice Boltzmann models

We present a unified approach of several boundary conditions for lattice Boltzmann models. Its general framework is a generalization of previously introduced schemes such as the bounce-back rule, linear or quadratic interpolations, etc. The objectives are two fold: first to give theoretical tools to study the existing boundary conditions and their corresponding accuracy; secondly to design formally third- order accurate boundary conditions for general flows. Using these boundary conditions, Couette and Poiseuille flows are exact solution of the lattice Boltzmann models for a Reynolds number $Re = 0$ (Stokes limit).
Numerical comparisons are given for Stokes flows in periodic arrays of spheres and cylinders, linear periodic array of cylinders between moving plates and for Navier-Stokes flows in periodic arrays of cylinders for $Re < 200$. These results show a significant improvement of the overall accuracy when using the linear interpolations instead of the bounce-back reflection (up to an order of magnitude on the hydrodynamics fields). Further improvement is achieved with the new multi-reflection boundary conditions, reaching a level of accuracy close to the quasi-analytical reference solutions, even for rather modest grid resolutions and few points in the narrowest channels. More important, the pressure and velocity fields in the vicinity of the obstacles are much smoother with multi-reflection than with the other boundary conditions.
Finally the good stability of these schemes is highlighted by some simulations of moving obstacles: a cylinder between flat walls and a sphere in a cylinder.
Keywords: lattice Boltzmann equation, boudary condistions, bounce-back rule, Navier-Stokes equation
(72 pages, 2002)

39. R. Korn

Elementare Finanzmathematik

Im Rahmen dieser Arbeit soll eine elementar gehaltene Einführung in die Aufgabenstellungen und Prinzipien der modernen Finanzmathematik gegeben werden. Insbesondere werden die Grundlagen der Modellierung von Aktienkursen, der Bewertung von Optionen und der Portfolio-Optimierung vorgestellt. Natürlich können die verwendeten Methoden und die entwickelte Theorie nicht in voller Allgemeinheit für den Schulunterricht verwendet werden, doch sollen einzelne Prinzipien so heraus gearbeitet werden, dass sie auch an einfachen Beispielen verstanden werden können.
Keywords: Finanzmathematik, Aktien, Optionen, Portfolio-Optimierung, Börse, Lehrerweiterbildung, Mathematikunterricht
(98 pages, 2002)

40. J. Kallrath, M. C. Müller, S. Nickel

Batch Presorting Problems: Models and Complexity Results

In this paper we consider short term storage systems. We analyze presorting strategies to improve the effiency of these storage systems. The presorting task is called Batch PreSorting Problem (BPSP). The BPSP is a variation of an assigment problem, i. e., it has an assigment problem kernel and some additional constraints. We present different types of these presorting problems, introduce mathematical programming formulations and prove the NP-completeness for one type of the BPSP. Experiments are carried out in order to compare the different model formulations and to investigate the behavior of these models.
Keywords: Complexity theory, Integer programming, Assigment, Logistics
(19 pages, 2002)

41. J. Linn

On the frame-invariant description of the phase space of the Folgar-Tucker equation

The Folgar-Tucker equation is used in flow simulations of fiber suspensions to predict fiber orientation depending on the local flow. In this paper, a complete, frame-invariant description of the phase space of this differential equation is presented for the first time.
Key words: fiber orientation, Folgar-Tucker equation, injection molding
(5 pages, 2003)

42. T. Hanne, S. Nickel

A Multi-Objective Evolutionary Algorithm for Scheduling and Inspection Planning in Software Development Projects

In this article, we consider the problem of planning inspections and other tasks within a software development (SD) project with respect to the objectives quality (no. of defects), project duration, and costs. Based on a discrete-event simulation model of SD processes comprising the phases coding, inspection, test, and rework, we present a simplified formulation of the problem as a multiobjective optimization problem. For solving the problem (i. e. finding an approximation of the efficient set) we develop a multiobjective evolutionary algorithm. Details of the algorithm are discussed as well as results of its application to sample problems.
Key words: multiple objective programming, project management and scheduling, software development, evolutionary algorithms, efficient set
(29 pages, 2003)

43. T. Bortfeld , K.-H. Küfer, M. Monz, A. Scherrer, C. Thieke, H. Trinkaus

Intensity-Modulated Radiotherapy - A Large Scale Multi-Criteria Programming Problem -

Radiation therapy planning is always a tight rope walk between dangerous insufficient dose in the target volume and life threatening overdosing of organs at risk. Finding ideal balances between these inherently contradictory goals challenges dosimetrists and physicians in their daily practice. Today's planning systems are typically based on a single evaluation function that measures the quality of a radiation treatment plan. Unfortunately, such a one dimensional approach can-

not satisfactorily map the different backgrounds of physicians and the patient dependent necessities. So, too often a time consuming iteration process between evaluation of dose distribution and redefinition of the evaluation function is needed.
In this paper we propose a generic multi-criteria approach based on Pareto's solution concept. For each entity of interest - target volume or organ at risk a structure dependent evaluation function is defined measuring deviations from ideal doses that are calculated from statistical functions. A reasonable bunch of clinically meaningful Pareto optimal solutions are stored in a data base, which can be interactively searched by physicians. The system guarantees dynamical planning as well as the discussion of tradeoffs between different entities.
Mathematically, we model the upcoming inverse problem as a multi-criteria linear programming problem. Because of the large scale nature of the problem it is not possible to solve the problem in a 3D-setting without adaptive reduction by appropriate approximation schemes.
Our approach is twofold: First, the discretization of the continuous problem is based on an adaptive hierarchical clustering process which is used for a local refinement of constraints during the optimization procedure. Second, the set of Pareto optimal solutions is approximated by an adaptive grid of representatives that are found by a hybrid process of calculating extreme compromises and interpolation methods.
Keywords: multiple criteria optimization, representative systems of Pareto solutions, adaptive triangulation, clustering and disaggregation techniques, visualization of Pareto solutions, medical physics, external beam radiotherapy planning, intensity modulated radiotherapy
(31 pages, 2003)

44. T. Halfmann, T. Wichmann

Overview of Symbolic Methods in Industrial Analog Circuit Design

Industrial analog circuits are usually designed using numerical simulation tools. To obtain a deeper circuit understanding, symbolic analysis techniques can additionally be applied. Approximation methods which reduce the complexity of symbolic expressions are needed in order to handle industrial-sized problems. This paper will give an overview to the field of symbolic analog circuit analysis. Starting with a motivation, the state-of-the-art simplification algorithms for linear as well as for nonlinear circuits are presented. The basic ideas behind the different techniques are described, whereas the technical details can be found in the cited references. Finally, the application of linear and nonlinear symbolic analysis will be shown on two example circuits.
Keywords: CAD, automated analog circuit design, symbolic analysis, computer algebra, behavioral modeling, system simulation, circuit sizing, macro modeling, differential-algebraic equations, index
(17 pages, 2003)

45. S. E. Mikhailov, J. Orlik

Asymptotic Homogenisation in Strength and Fatigue Durability Analysis of Composites

Asymptotic homogenisation technique and two-scale convergence is used for analysis of macro-strength and fatigue durability of composites with a periodic structure under cyclic loading. The linear damage accumulation rule is employed in the phenomenological micro-durability conditions (for each component of the composite) under varying cyclic loading. Both local and non-local strength and durability conditions are analysed. The strong convergence of the strength and fatigue damage measure as the structure period tends to zero is proved and their limiting values are estimated.
Keywords: multiscale structures, asymptotic homogenization, strength, fatigue, singularity, non-local conditions
(14 pages, 2003)

46. P. Domínguez-Marín, P. Hansen, N. Mladenović, S. Nickel

Heuristic Procedures for Solving the Discrete Ordered Median Problem

We present two heuristic methods for solving the Discrete Ordered Median Problem (DOMP), for which no such approaches have been developed so far. The DOMP generalizes classical discrete facility location problems, such as the p-median, p-center and Uncapacitated Facility Location problems. The first procedure proposed in this paper is based on a genetic algorithm developed by Moreno Vega [MV96] for p-median and p-center problems. Additionally, a second heuristic approach based on the Variable Neighborhood Search metaheuristic (VNS) proposed by Hansen & Mladenovic [HM97] for the p-median problem is described. An extensive numerical study is presented to show the efficiency of both heuristics and compare them.
Keywords: genetic algorithms, variable neighborhood search, discrete facility location
(31 pages, 2003)

47. N. Boland, P. Domínguez-Marín, S. Nickel, J. Puerto

Exact Procedures for Solving the Discrete Ordered Median Problem

The Discrete Ordered Median Problem (DOMP) generalizes classical discrete location problems, such as the N-median, N-center and Uncapacitated Facility Location problems. It was introduced by Nickel [16], who formulated it as both a nonlinear and a linear integer program. We propose an alternative integer linear programming formulation for the DOMP, discuss relationships between both integer linear programming formulations, and show how properties of optimal solutions can be used to strengthen these formulations. Moreover, we present a specific branch and bound procedure to solve the DOMP more efficiently. We test the integer linear programming formulations and this branch and bound method computationally on randomly generated test problems.
Keywords: discrete location, Integer programming
(41 pages, 2003)

48. S. Feldmann, P. Lang

Padé-like reduction of stable discrete linear systems preserving their stability

A new stability preserving model reduction algorithm for discrete linear SISO-systems based on their impulse response is proposed. Similar to the Padé approximation, an equation system for the Markov parameters involving the Hankel matrix is considered, that here however is chosen to be of very high dimension. Although this equation system therefore in general cannot be solved exactly, it is proved that the approximate solution, computed via the Moore-Penrose inverse, gives rise to a stability preserving reduction scheme, a property that cannot be guaranteed for the Padé approach. Furthermore, the proposed algorithm is compared to another stability preserving reduction approach, namely the balanced truncation method, showing comparable performance of the reduced systems. The balanced truncation method however starts from a state space description of the systems and in general is expected to be more computational demanding.
Keywords: Discrete linear systems, model reduction, stability, Hankel matrix, Stein equation
(16 pages, 2003)

49. J. Kallrath, S. Nickel

A Polynomial Case of the Batch Presorting Problem

This paper presents new theoretical results for a special case of the batch presorting problem (BPSP). We will show tht this case can be solved in polynomial time. Offline and online algorithms are presented for solving the BPSP. Competetive analysis is used for comparing the algorithms.
Keywords: batch presorting problem, online optimization, competetive analysis, polynomial algorithms, logistics
(17 pages, 2003)

50. T. Hanne, H. L. Trinkaus

knowCube for MCDM – Visual and Interactive Support for Multicriteria Decision Making

In this paper, we present a novel multicriteria decision support system (MCDSS), called knowCube, consisting of components for knowledge organization, generation, and navigation. Knowledge organization rests upon a database for managing qualitative and quantitative criteria, together with add-on information. Knowledge generation serves filling the database via e. g. identification, optimization, classification or simulation. For "finding needles in haycocks", the knowledge navigation component supports graphical database retrieval and interactive, goal-oriented problem solving. Navigation "helpers" are, for instance, cascading criteria aggregations, modifiable metrics, ergonomic interfaces, and customizable visualizations. Examples from real-life projects, e.g. in industrial engineering and in the life sciences, illustrate the application of our MCDSS.
Key words: Multicriteria decision making, knowledge management, decision support systems, visual interfaces, interactive navigation, real-life applications.
(26 pages, 2003)

51. O. Iliev, V. Laptev

On Numerical Simulation of Flow Through Oil Filters

This paper concerns numerical simulation of flow through oil filters. Oil filters consist of filter housing (filter box), and a porous filtering medium, which completely separates the inlet from the outlet. We discuss mathematical models, describing coupled flows in the pure liquid subregions and in the porous filter media, as well as interface conditions between them. Further, we reformulate the problem in fictitious regions method manner, and discuss peculiarities of the numerical algorithm in solving the coupled system. Next, we show numerical results, validating the model and the

algorithm. Finally, we present results from simulation of 3-D oil flow through a real car filter.
Keywords: oil filters, coupled flow in plain and porous media, Navier-Stokes, Brinkman, numerical simulation
(8 pages, 2003)

52. W. Dörfler, O. Iliev, D. Stoyanov, D. Vassileva

On a Multigrid Adaptive Refinement Solver for Saturated Non-Newtonian Flow in Porous Media

A multigrid adaptive refinement algorithm for non-Newtonian flow in porous media is presented. The saturated flow of a non-Newtonian fluid is described by the continuity equation and the generalized Darcy law. The resulting second order nonlinear elliptic equation is discretized by a finite volume method on a cell-centered grid. A nonlinear full-multigrid, full-approximation-storage algorithm is implemented. As a smoother, a single grid solver based on Picard linearization and Gauss-Seidel relaxation is used. Further, a local refinement multigrid algorithm on a composite grid is developed. A residual based error indicator is used in the adaptive refinement criterion. A special implementation approach is used, which allows us to perform unstructured local refinement in conjunction with the finite volume discretization. Several results from numerical experiments are presented in order to examine the performance of the solver.
Keywords: Nonlinear multigrid, adaptive refinement, non-Newtonian flow in porous media
(17 pages, 2003)

53. S. Kruse

On the Pricing of Forward Starting Options under Stochastic Volatility

We consider the problem of pricing European forward starting options in the presence of stochastic volatility. By performing a change of measure using the asset price at the time of strike determination as a numeraire, we derive a closed-form solution based on Heston's model of stochastic volatility.
Keywords: Option pricing, forward starting options, Heston model, stochastic volatility, cliquet options
(11 pages, 2003)

54. O. Iliev, D. Stoyanov

Multigrid – adaptive local refinement solver for incompressible flows

A non-linear multigrid solver for incompressible Navier-Stokes equations, exploiting finite volume discretization of the equations, is extended by adaptive local refinement. The multigrid is the outer iterative cycle, while the SIMPLE algorithm is used as a smoothing procedure. Error indicators are used to define the refinement subdomain. A special implementation approach is used, which allows to perform unstructured local refinement in conjunction with the finite volume discretization. The multigrid - adaptive local refinement algorithm is tested on 2D Poisson equation and further is applied to a lid-driven flows in a cavity (2D and 3D case), comparing the results with bench-mark data. The software design principles of the solver are also discussed.
Keywords: Navier-Stokes equations, incompressible flow, projection-type splitting, SIMPLE, multigrid methods, adaptive local refinement, lid-driven flow in a cavity
(37 pages, 2003)

55. V. Starikovicius

The multiphase flow and heat transfer in porous media

In first part of this work, summaries of traditional Multiphase Flow Model and more recent Multiphase Mixture Model are presented. Attention is being paid to attempts include various heterogeneous aspects into models. In second part, MMM based differential model for two-phase immiscible flow in porous media is considered. A numerical scheme based on the sequential solution procedure and control volume based finite difference schemes for the pressure and saturation-conservation equations is developed. A computer simulator is built, which exploits object-oriented programming techniques. Numerical result for several test problems are reported.
Keywords: Two-phase flow in porous media, various formulations, global pressure, multiphase mixture model, numerical simulation
(30 pages, 2003)

56. P. Lang, A. Sarishvili, A. Wirsen

Blocked neural networks for knowledge extraction in the software development process

One of the main goals of an organization developing software is to increase the quality of the software while at the same time to decrease the costs and the duration of the development process. To achieve this, various decisions e.ecting this goal before and during the development process have to be made by the managers. One appropriate tool for decision support are simulation models of the software life cycle, which also help to understand the dynamics of the software development process. Building up a simulation model requires a mathematical description of the interactions between di.erent objects involved in the development process. Based on experimental data, techniques from the .eld of knowledge discovery can be used to quantify these interactions and to generate new process knowledge based on the analysis of the determined relationships. In this paper blocked neuronal networks and related relevance measures will be presented as an appropriate tool for quanti.cation and validation of qualitatively known dependencies in the software development process.
Keywords: Blocked Neural Networks, Nonlinear Regression, Knowledge Extraction, Code Inspection
(21 pages, 2003)

Status quo: November 2003