18 A TENDERING PROCESS MODEL FOR AN APPLICATION OF RELIABILITY ANALYSIS

André Bester
University of Stellenbosch, abester@ing.sun.ac.za

Wolfgang Huhnt
Bauhaus-Universität Weimar, wolfgang.huhnt@informatik.uni-weimar.de

Abstract

Construction data is a useful source of information for construction project management. The sources of data for various aspects of construction management is, in addition to the current project information and the engineer's professional experience, also historical data from previous projects. An example of the use of historical cost data is in the application of the Limit State Cost Function (LSCF) model to perform construction project risk modelling.

The LSCF-method was developed for assisting the consulting civil engineer to make a quantifiable judgement in the selection of a project tender. It is based on the reliability of estimating the outcome of the cost for a specific tender in terms of the risks involved in a project, by using construction cost data.

The way in which this data have been captured has shown deficiencies which limits the utility of it if the scope of the data captured and the processes of both capturing and applying the results are not carefully planned. The application of the data can therefore not be seen without the process in which it is generated.

The process is characterised by the generation and processing of data by the co-operation of several groups of persons at the different stages of the project. These groups of persons execute different tasks, for which appropriate subsets of information at different levels of detail are needed.

The paper is focused on the specific requirements that have to be supported to model such processes. Tools are presented to support the specification of the process model. These tools support the investigation of the model with respect to completeness and correctness. In addition, the tools support the implementation of an information system that is to be used for tasks during the tender process.

Keywords: *tender, process, reliability, sets, graphs.*

INTRODUCTION

Process Modelling

The question how process models can be used in a beneficial way during the execution of projects in construction companies is still very much in dispute. Many investigative studies in process modelling have been undertaken in other fields of engineering tasks. For instance, process models are used in the software industry to monitor information systems or in the motor industry to analyse, optimise and simulate industrial processes.

Experience, in the use of process models, have however shown that these models can not be directly adapted to the construction industry. Construction projects are characterised by some specific peculiarities. In general, a project is executed only once and therefore mass production is not applicable. Changes to the project are also regularly been made during the execution of a project and therefore the modelled process changes.

It is necessary though, to describe the sequences of tasks, including the dependencies between them. Construction projects are executed by several persons who are usually not working at the same location so that a co-ordination of all tasks is necessary in addition to the execution of the tasks themselves.

The approach described in this paper focuses on two aspects in the use of process models in construction projects: the process model is used to determine whether it describes a complete and correct (workable) sequence of tasks; it is used to simplify the implementation of software to support the execution of these tasks.

The considerations are based on a specific method for the specification of process models. Only a manageable number of object types is introduced to specify the process model. These object types allow for the specification of suitable process models without additional and unnecessary effort. The level of detail can not be changed in the model so that the model becomes clear and comprehensible.

The advantage of the use of the method for the specification of the process model is shown in the field of tender processes. It includes investigations with respect to correctness and completeness and the implementation of software to support the execution of the work.

Reliability Analysis

At the tender evaluation stage, when the expected cost of construction of the project is known for each tender, a decision must be made to whom the execution of the construction project is awarded. This decision is made either by or in conjunction with the client and consulting engineer. Usually the project is awarded to the lowest bidder, if this bidder proves to be financially sound and has the necessary experience.

Unexpected costs due to risks in a project is usually not included in the evaluation of the tendered bids, except empirically by adding a percentage of the total tender or a percentage on specific parts of the tender to the total, as a contingency. This method does not take into account the influence the project risks will have on the total costs of the project.

The problem that exists is two-folded. Firstly, the analysis techniques employed are at best sensitivity analysis of single variables while the rest of the variables which can also influence the total cost are treated deterministically (single fixed values). Probabilistic analysis where variables have a range of values and probabilities of occurrence solves this problem, e.g. Limit State Cost Function [1], that is based on reliability analysis [2].

In the application of this probabilistic technique information about the statistical distributions and its parameters are needed. The two sources of information are the experience of the engineer in guessing the distributions and its parameters for the variables or, the more important one, historical data.

This then gives rise to the second part of the problem, unfortunately the quality and availability of historical data, which should be used, is poor. Two of the most valuable information sources of historical data, the Tender Report and the Final Report of projects are never captured and stored for later analysis. In the few cases where this is done the process suffers of multiple recapturing of almost identical data, data losses and inconsistencies.

Solving the second problem requires the improvement of the data flow process that takes place from when the data is created until it is needed at a later stage. The type and amount of information captured is dependent on the usage of it. In order to improve both the analysis techniques and the data capturing process it is therefore necessary to realise that the two is interdependent from one another.

The accuracy of the Limit State Cost Function analysis technique can greatly be improved if the distributions and their parameters could be calculated better. This can be done by taking into account the effect the different risks of a project have on the characteristics of the variables. This type of information can be found if one analysis the data generated in different projects. It is therefore necessary to do investigations into the data flow process and the process model in which the data is generated as to improve the capturing of appropriate data.

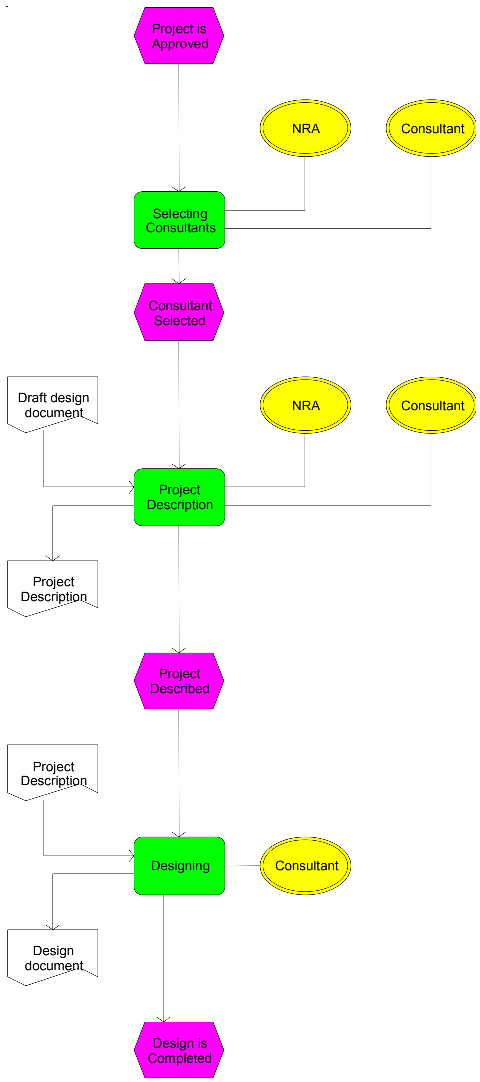

Figure 1: A Part of the Tendering Process Model

SPECIFICATION OF THE PROCESS MODEL

For the specification of the process model, four object types are introduced: tasks, events, groups of persons (groups) and sets of data (data). Tasks are used to specify the work that is to be executed. The execution of a task needs time. Events are used to specify states. They describe a point of time in the sequence of the process. Groups are used to specify the role of the persons involved in the process. Data is used for the execution of the tasks.

Two types of relations can be modelled between tasks and events. The relation between a task and an event denotes the event that can occur after the execution of the task, the relation between an event and a task denotes the task that can occur as a consequence of the event. These two relations are used to specify the sequence of the tasks.

A relation can be modelled to denote the groups of persons who are involved in the execution of a task. Between a task and a set of data, four relations can be modelled. These relations denote whether a data in the set is created, written (changes), read or deleted during the execution of the task.

Figure 1 shows a part of the tendering process model. This part was modelled using an existing tool set [3]. The model only makes use of the types of objects and the relations as described above whereas tasks are visualised by rectangles with rounded edges, events by hexagons, persons by ellipses and documents by truncated rectangles.

THEORETICAL BACKGROUND

The theoretical background is based on the set theory [4]. Four sets are created during the specification of the process model where each set consists of objects that are of the same type. As an example, a few of the elements of the tender process model are shown for each set:

- Task = {Define Project; Conceptual Design; Designing; Cost Calculation }
- Event = {Project Defined; Conceptual Design Completed; Design Approved; Cost Calculated }
- Group = {Client; Consultant; Contractor; Sub-Contractor 1; Sub-Contractor 2}
- Data = {Codes; Project Specifications; Items; Tenders}

On the base of these sets the relations are modelled:

- The sequence between tasks and events: TE = {(x,y) x ∈ Task, y ∈ Event | leads to}
- The sequence between events and tasks: ET = {(x,y) x ∈ Event, y ∈ Task | leads to}
- A task creates data: TC = {(x,y) x ∈ Task, y ∈ Data | creates}
- A task writes data: TW = {(x,y) x ∈ Task, y ∈ Data | writes}
- A task reads data: TR = {(x,y) x ∈ Task, y ∈ Data | reads}
- A task deletes data: TD = {(x,y) x ∈ Task, y ∈ Data | deletes}
- A group executes a task: TG = {(x,y) x ∈ Task, y ∈ Group | involves}

Some properties of the relations are known a priori. The relation TG is symmetric. The sets Task and Event and the relations TE and ET are a bipartite graph.

The specification on the base of the set theory allows for the execution of algorithms to determine properties of the sets and the relations. In addition, the sets and the relations can be mapped onto graphs so those routes can be determined on the base of the Boolean algebra [5]. These algorithms are used to investigate the process model as described in the next section.

INVESTIGATIONS

Several questions can be asked that will address tests for correctness and completeness of a process model. The questions are grouped by the sets involved to produce an answer. The algorithms to answer these questions are based on the set and the graph theory. As an example, results are presented which are based on the tender process model. These results have been determined in a pilot implementation.

Task and Group

The relation *TG*, between the set *Task* and the set *Group* is symmetric and can be described by the adjacency matrix *M*, where:

$$m_{ij} = \begin{cases} 1, & if\,(task_i, group_j) \in TG. \\ 0, & otherwise. \end{cases}$$

Question: How many groups are involved in a task?
Algorithm: The sum of each row in *M* can be taken, giving the number of *Group*s involved in a *Task*. Results are shown in figure 2.

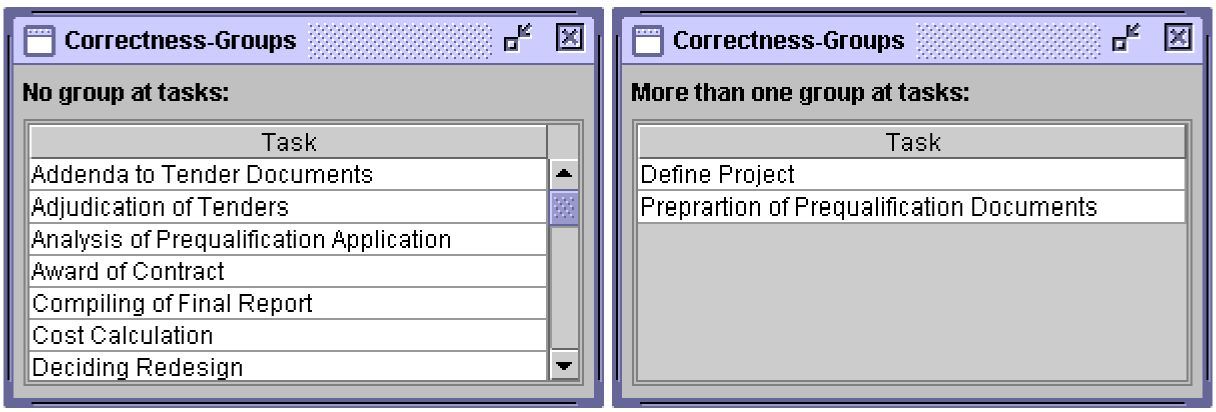

Figure 2: Correctness of the Relation between Groups and Tasks

Question: In what tasks is a group involved?
Algorithm: Search down the column of the given *Group* in *M* and find the *Task*s in the corresponding rows indicated by 1's in the column. Results are shown in figure 3.

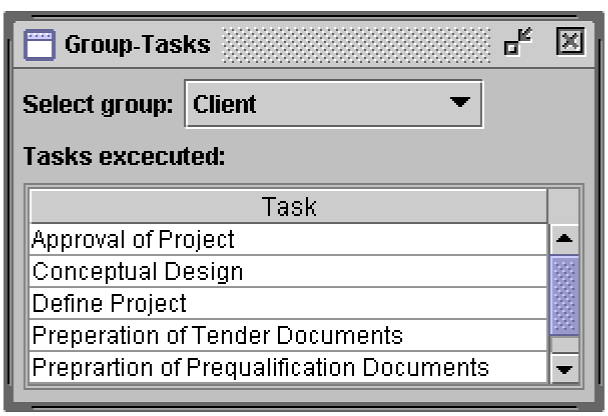

Figure 3: Groups and Tasks

Question: Who is involved in a task?
Algorithm: Search down the row of the given *Task* in *M* and find the *Group*s in the corresponding columns indicated by 1's in the row. Results are shown in figure 4.

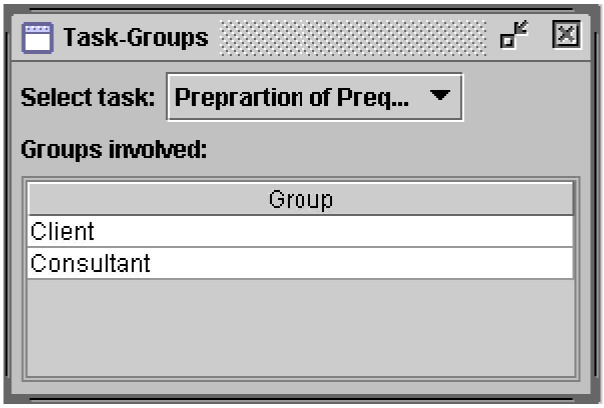

Figure 4: Tasks and Groups

Task and Data

There exist four relations between *Task* and *Data*: *create*, *write*, *read* and *delete*. When one wants to analyse the process for correctness of data access, it has to be recognised that databases distinguish between *read*-access and *write*-access only. *Create*, *write* and *delete* qualifies as *write*-access to the database and *read* as *read*-access when generating the database access on the base of the process model.

To investigate the question whether data is available when needed, the data have to be created before they can be written (changed), read or deleted. For this purpose, the relation *create* has to appear in the model before one of the relations *write*, *read* and *delete* is specified. The relation *create* is regarded as *write*-access whereas the relations *write*, *read* and *delete* as *read*-access.

In order to distinguish between these two access models it is necessary to define two incidence matrices, one for the process data access, matrix *P*, and one for the standard data access, matrix *D*:

$$p_{ij} = \begin{cases} 1, & \text{if } task_i \text{ writes } data_j \\ -1, & \text{if } task_i \text{ reads } data_j \\ 0 & \text{otherwise.} \end{cases}$$

and

$$d_{ij} = \begin{cases} 1, & \text{if } task_i \text{ writes } data_j \\ -1, & \text{if } task_i \text{ reads } data_j \\ 0 & \text{otherwise.} \end{cases}$$

Question: Which tasks use data and what is the type of access in the process or data model?
Algorithm: Search through the column corresponding to the selected *Data* in matrix *P* or *D* and find the rows corresponding to *Task* as well as access which depends on the value p_{ij} or d_{ij}. Results are shown in figure 5.
Question: Which data is used by a task and what is the type of access in the process or data model?

Algorithm: Search through the row corresponding to the selected *Task* in matrix *P* or *D* and find the columns corresponding to *Data* as well as access which depends on the value p_{ij} or d_{ij}. Results are shown in figure 6.
Question: How many data are involved at a task?
Algorithm: The sum of each row in either *P* or *D* can be taken, giving the number of *Group*s involved in a *Task*. Results are shown in figure 7.

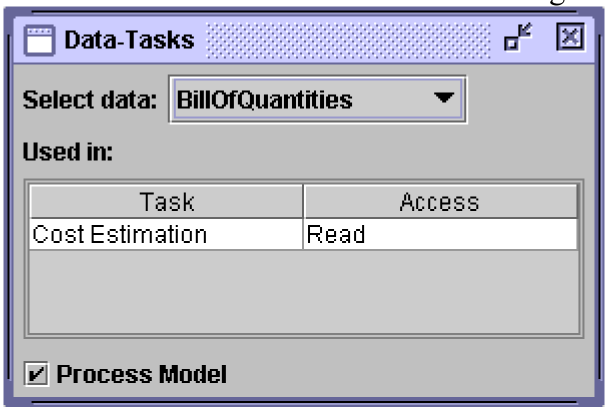

Figure 5: Use of Data in Tasks

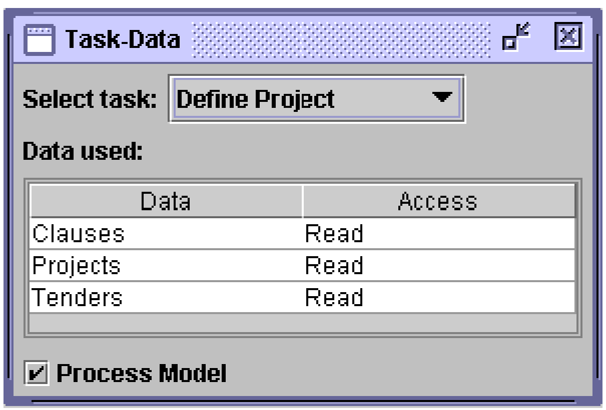

Figure 6: Access from Tasks to Data

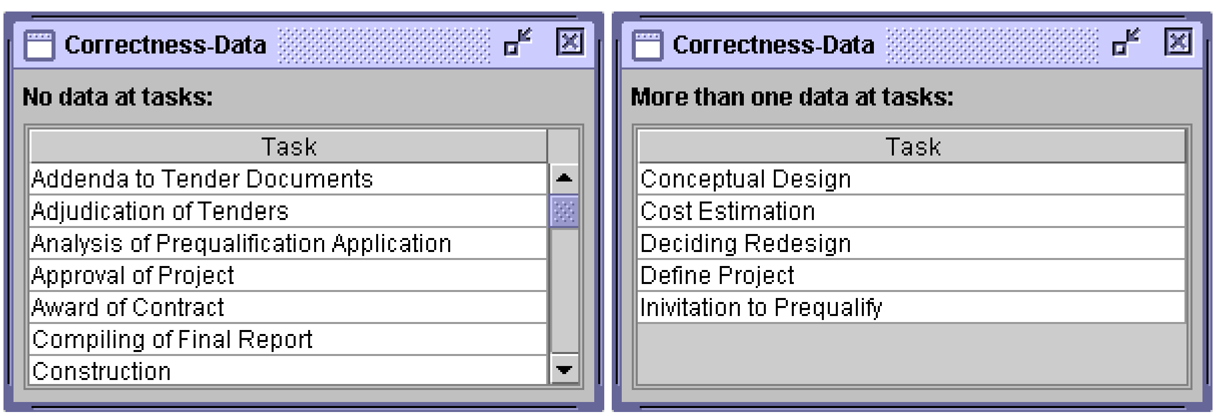

Figure 7: Correctness of the Relation between Data and Tasks

The following question concerning the sets *Task* and *Data* can also be formulated:

- Is the sequence of write and read correctly defined?
- Which tasks have changed this data or read it before this task?
- Which task will change this data or read it after this task?
- Which tasks are affected by erroneous data?

In order to answer these questions the process model has to include how long it takes for a process to be completed before the next one starts. As mentioned previously the process model is mathematically modelled with a directed graph. In order to be able to include time duration a weighted graph might be considered however this question is beyond the scope of this paper and is therefore not considered.

Data and Group

The relation *TG*, between the set *Task* and the set *Group* is symmetric and can be described by the adjacency matrix:

$$m_{ij} = \begin{cases} 1, & if\,(task_i, group_j) \in TG. \\ 0, & otherwise. \end{cases}$$

In order to again distinguish between the two access models it is necessary to define two incidence matrices, one for the process data access, matrix *P* and one for the standard data access, matrix *D*:

$$p_{ij} = \begin{cases} 1, & if\ task_i\ writes\ data_j \\ -1, & if\ task_i\ reads\ data_j \\ 0 & otherwise. \end{cases}$$

and

$$d_{ij} = \begin{cases} 1, & if\ task_i\ writes\ data_j \\ -1, & if\ task_i\ reads\ data_j \\ 0 & otherwise. \end{cases}$$

Also define the following two matrices *K* and *L*, where *K* gives the relationship that exist between *Group* and *Data* for the process model access and *L* for the same relationship, but for the data model access:

$$k_{ij} = \begin{cases} 1, & if\ group_i\ writes\ data_j \\ -1, & if\ group_i\ reads\ data_j \\ 2, & if\ group_i\ reads\ and\ writes\ data_j \\ 0 & otherwise. \end{cases}$$

and

$$l_{ij} = \begin{cases} 1, & if\ group_i\ writes\ data_j \\ -1, & if\ group_i\ reads\ data_j \\ 2, & if\ group_i\ reads\ and\ writes\ data_j \\ 0 & otherwise. \end{cases}$$

The calculation of these two matrices is based on matrix multiplication between the transpose of matrix M^T and matrix K or L. The only difference is at the point where the sum of the products should be calculated, in the straightforward algorithm the matrix product C of two n×n matrices A and B are defined as follows:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

where as for this application it is defined as:

$$c_{ij} = \begin{cases} 1, & \textit{if } \exists a_{ik} b_{kj} = 1 \textit{ and } \forall a_{ik} b_{kj} \neq -1 \textit{ for } k = 1..n \\ -1, & \textit{if } \exists a_{ik} b_{kj} = -1 \textit{ and } \forall a_{ik} b_{kj} \neq 1 \textit{ for } k = 1..n \\ 2, & \textit{if } \exists a_{ik} b_{kj} = -1 \textit{ and } \exists a_{ik} b_{kj} = 1 \textit{ for } k = 1..n \\ 0 & \textit{otherwise.} \end{cases}$$

Question: Who uses data and what is the type of access in the process or data model?
Algorithm: Search through the column corresponding to the selected *Data* in matrix K or L and find the rows corresponding to *Group* as well as access which depends on the value k_{ij} or l_{ij}. Results are shown in figure 8.

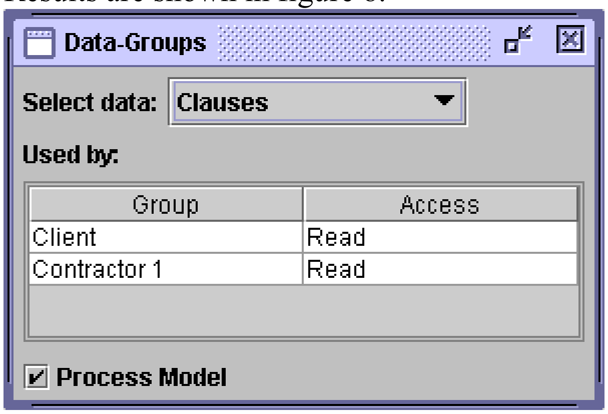

Figure 8: Use of Data by Groups of Persons

Question: Which data is used by a group and what is the type of access in the process or data model?
Algorithm: Search through the row corresponding to the selected *Group* in matrix K or L and find the columns corresponding to *Data* as well as access which depends on the value k_{ij} or l_{ij}. Results are shown in figure 9.

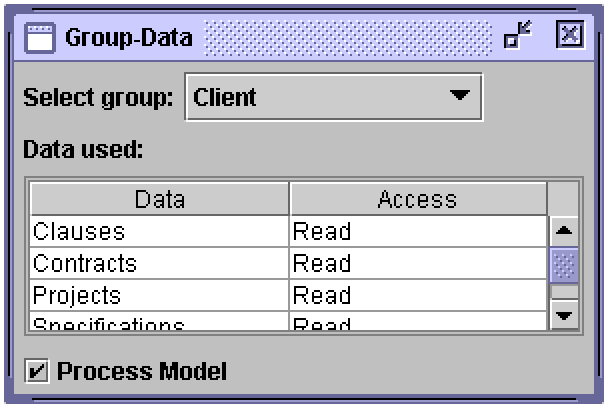

Figure 9: Access from Groups of Persons to Data

The following question concerning the sets *Task* and *Data* can also be formulated:

- Which groups have used the data before this task and what is the access?
- Which groups will use this data after this group and what is the access?
- Which groups are affected by errors in data of a previous task?

To be able to answer these questions it is again necessary to consider the duration of time before completion, therefore it is not further discussed as it is beyond the scope of this paper.

Tasks and Events

In the process model, it can happen that an event or task leads to two or more tasks or events respectively. There are two reasons for this. Processes are executed in parallel to one another and will join at a later stage or end while another continues. In graph theory this is a so-called directed path. In the second case it might happen that a process returns to previous events or task. In this case we have a cycle.

Both these cases are important to recognise in a process model because one would like to know where these situations develop. This is necessary in order for one to be able to answer questions with regard to whether processes actually take place in parallel or whether a cycle exists; who is involved and what data is used or created; who is responsible to make a decision if a cycle should continue or not. The identification of these two cases can be done as follows.

Generate an incidence matrix by using the directed connectivity of *Tasks* and *Events* with the following properties:

$$a_{ij} = \begin{cases} 1, & \text{if the } j\text{th edge is incident inward on the } i\text{th vertices} \\ -1, & \text{if the } j\text{th edge is incident outward on the } i\text{th vertices} \\ 0 & \text{otherwise} \end{cases}$$

Question: Are there any cycles or paths?
Algorithm: Run a depth-first search algorithm, which has been modified to test for strong connectivity, on matrix A.

The branches, which form from the depth-first search algorithm, can either be cycles or paths.

The strongly connected components (branches) which are identified by the strong connectivity test are the cycles while the other branches are the paths. Results are shown in figure 10.

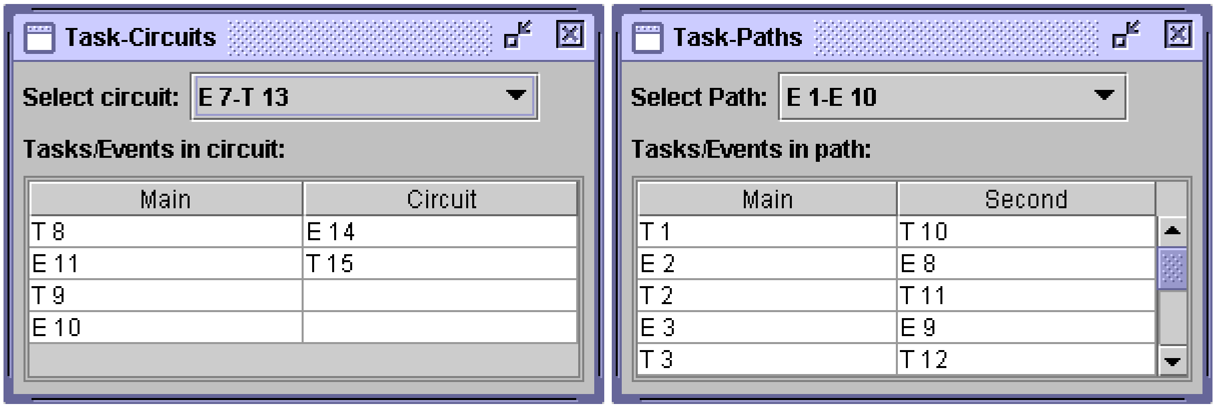

Figure 10: Cycles and Paths

IMPLEMENTATION

The process model can be regarded as a road map for the implementation of an information system to support the required tasks. It allows for investigations concerning correctness as describes above. Therefore, it is useful for the target specification in the process of software development, and it can be helpful for the communication between the software developer and the client.

In addition, the process model can simplify the effort of implementation: the specification of the data sets can be used to generate a database schema. Each data set consists of elements that are of the same type. For the specification of object types base sets are introduced: String, Integer and Double. An object type consists of ordered pairs whereas the first component denotes the semantic and the second component the set where the specific value is taken from (data type). In addition, relations and set systems can be specified.

The mathematical background of this specification allows for both, the generation of a schema for a relational database and the generation of a schema for an object oriented database. The pilot implementation of the tools to specify and investigate the process model supports only relational databases [6]. Figure 2 shows an overview on the database schema, which was generated on the base of the process model. This database schema was used for a pilot implementation to support the tendering process.

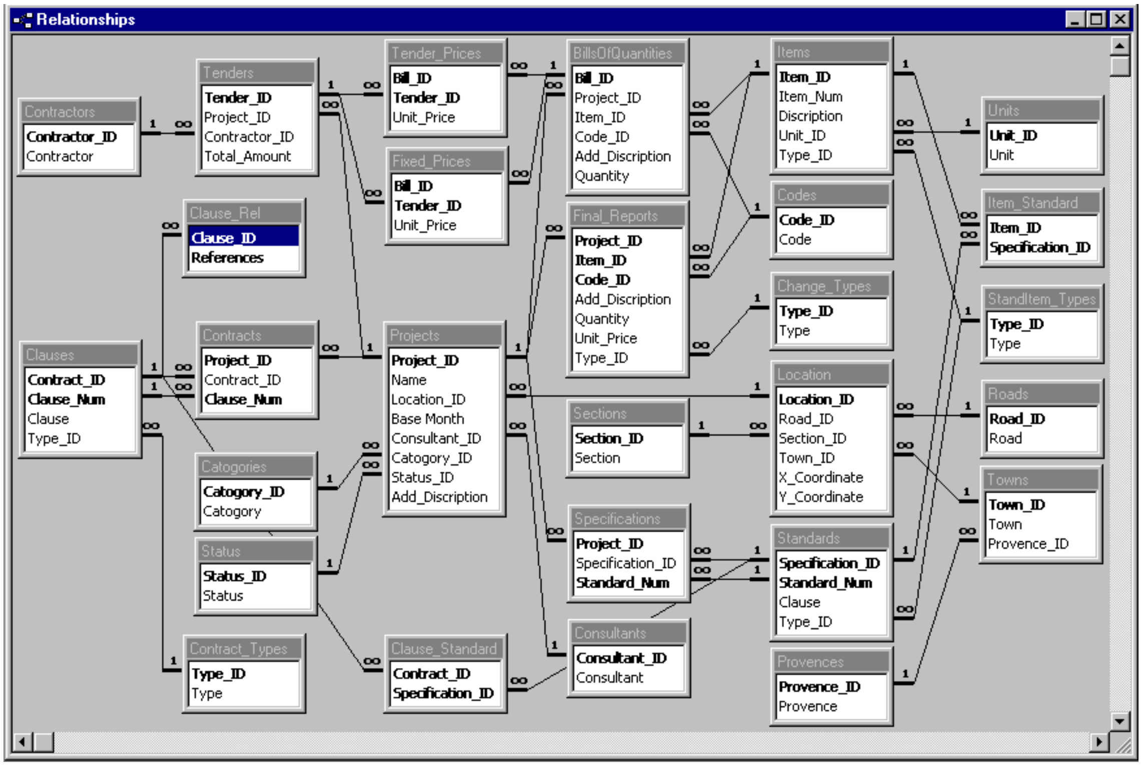

Figure 11: Generated Database Schema

A third aspect in the use of the process model is given in the phase when the software is used by the client. On the base of the process model, the running system can be monitored. Tables in a data base can be addressed. Data in these tables can be visualised to show the actual state of the system.

CONCLUSIONS

It was found that the questions on correctness and completeness could be answered on the base of a process model. This was achieved by introducing a manageable number of object types for the specification of process models and by using the set theory as the theoretical base for the process models. Based on the set theory algorithm can be used to answer questions which are import to evaluate process models. The algorithms described in the paper are simple well-known algorithms in the set and graph theory. It was shown that these algorithms could be used in a beneficial way. In addition, it was shown that the mathematical description on the base of the set theory could be used to reduce the effort for the implementation of an information system.

The main consideration was to illustrate a way where process models in civil and building engineering can be used in a beneficial way without a great effort of specification. For this purpose, only a simple mathematical background was used. Enhancements can be introduced. For instance, the duration of each task can be specified. However, each enhancement needs an additional effort of specification and an expanded mathematical background. In the case of the specification of the duration of the tasks, weighted graphs can be used to simulate the process. The question is whether the results of a process simulation justify the additional effort of specification.

Further investigations are necessary to answer this question in a more generalised way. It is necessary to find the balance between the effort of specification and the advantage of process models in civil and building engineering.

LITERATURE

Ker-Fox, G.M., (1998), *The Limit State Cost Function as a Risk Management Tool for Construction Projects*, Master Thesis, University of Stellenbosch Printers, Stellenbosch.

Ang, A.H.S., Tand, W.H., (1975), *Probability Concept in Engineering Planning and Design Volume 2*, John Wiley & Sons, New York.

ARIS Toolset, Version 4.0, IDS Scheer AG, Germany.

Pahl, P.J., Damrath, R., (2000), *Mathematische Grundlagen der Ingenieurinformatik,* Springer Verlag, Berlin.

Swamy, M.N.S., Thulasiraman, K., (1992), *Graphs, Networks and Algorithms*, John Wiley & Sons, Inc., New York.

Date, C. J., (1995), *An Introduction to Database Systems*, Addison-Wesley Publishing Company, Reading.