# Errors in the SAP Reference Model

**Jan Mendling[1], Wil van der Aalst[2], Boudewijn van Dongen[2], and Eric Verbeek[2]**

[1]Vienna University of Economics and Business Administration,
Augasse 2-6, A-1180 Vienna, Austria
jan.mendling@wu-wien.ac.at

[2]Technische Universiteit Eindhoven,
Postbus 513, NL-5600 MB Eindhoven, The Netherlands
w.m.p.v.d.aalst@tm.tue.nl

The SAP Reference Model is a set of information models that is utilized to guide the configuration of SAP systems. A big part of these models are business process models represented in the Event-driven Process Chains (EPC) notation. These EPC models provide a easy to comprehend overview of SAP business functionality on an abstract level. As such, the SAP Reference Model is considered as an important tool to bridge the gap between business logic and information system implementation. In a recent research project, we found that there are several errors in the reference model: 5.6% of the more than 600 processes that we analyzed revealed errors.

## The SAP Reference Model

Since the middle of the 1990s the SAP R/3 Reference Model has been available in different versions to support the implementation and configuration of the SAP system. The reference model is not only included in the SAP system itself, but also shipped with the business process modeling tools ARIS of IDS Scheer or NetProcess of IntelliCorp. The reference model covers several modeling perspectives such as data and organization structure, but the main emphasis is on business processes represented as Event-driven Process Chains (EPC). EPCs are a process modeling notation offering a function element type to capture business activities, an event element type to describe the pre- and post-conditions of functions, and so-called connectors as routing elements. Connectors represent splits and joins in a process flow and define whether a logical AND, OR, or XOR is used as a rule for splitting or joining several branches. All these elements can be linked with control flow arcs and, as a constraint, functions and events have to alternate. (In other words, functions can not connect to functions or events to events. Both types must alternate on any path.) The SAP Reference Model contains almost 10,000 individual models with 3000 of them being EPCs. These models are used for several purposes. The EPC business processes can easily be compared with the existing processes of an individual company revealing which kind of customizations are needed to make the SAP system fit the company. Furthermore, the process models are an easy to comprehend representation of the business semantics making them a helpful tool to train new staff members. Finally, they can also be used as best practice recommendations. In this case, the SAP processes are used as a blueprint to modify the existing processes of a company. The SAP Reference Model has been used in several industry projects with at least one of these ideas in mind. Yet it is interesting to note that, although the reference model has been heavily used in practice, it contains several errors.

## Finding A Lower Bound For The Number Of Errors

From the 3000 EPCs in the reference model there are 604 with non-trivial control flow. We analyzed these models for errors stemming from incorrect combinations of connector elements such as deadlocks and livelocks. A deadlock describes a situation in a process model where e.g. a customer order remains waiting for an activity to complete that can never be executed. A simple pattern leading to a deadlock is an XOR split which is joined with an AND. A livelock is an infinite loop e.g. when a

customer order repeatedly is forwarded between two departments without ever being completed. A simple modeling error that leads to a livelock is loop with an AND split at its exit.

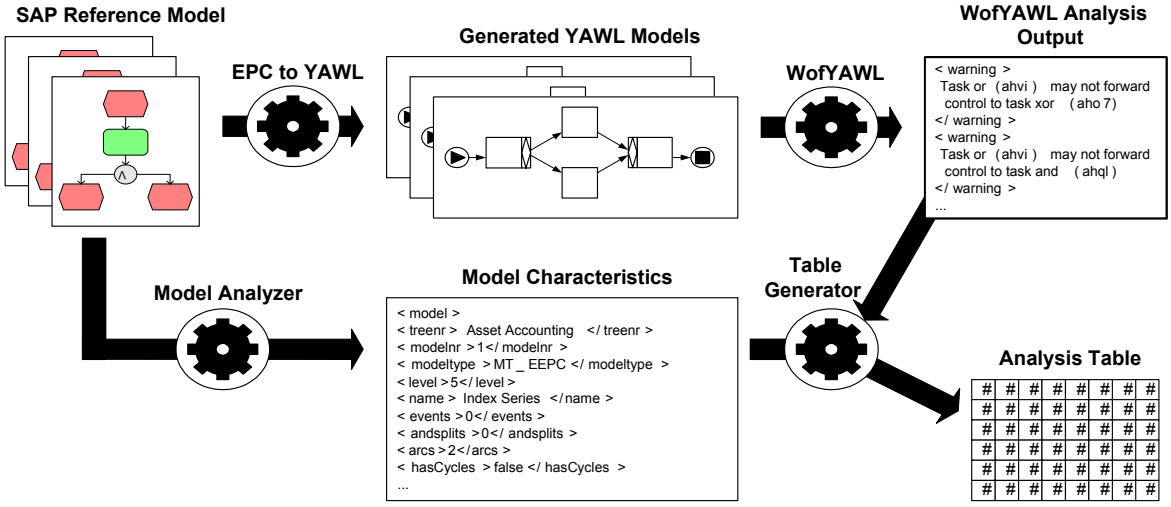


**Figure 1 An Overview of the Analysis Design**

Formal errors such as deadlocks and livelocks can be found with Petri nets analysis methods. Figure 1 gives an overview of our analysis design. First, we transformed the 604 EPCs to the workflow language YAWL. As EPCs offer a subset of the constructs that YAWL offers we could benefit from the analysis tool WofYAWL. This tool provides analysis functionality for YAWL models based on Petri nets. We furthermore extracted model characteristics such as the size of the EPC models or how often which connector was used and consolidated this data in an analysis table.

Based on this design we found that at least 34 of the 604 non-trivial EPCs contain errors, i.e. 5.6% of the models have errors. Table 1 gives an overview of the findings disaggregated by branches. Each branch represents a collection of processes that are supported by the respective SAP solution. In the reference model there are 29 branches each containing several process models. The procurement branch for example includes 444 models of which 37 are non-trivial EPCs. The last column indicates the relative amount of erroneous models that we identified using the automatic analysis described above. In this context it has to be noted that the 5.6% is only a lower bound for the number errors as our automatic analysis does not reveal all errors. Many errors require the interpretation of the corresponding model in order to be discovered. As an example consider the Procurement branch: using our automatic analysis we find only 2 errors. However, a slightly more refined analysis technique based on reductions and the explicit indication of valid sets start events already identifies 2 more errors in this branch (i.e., 4 errors). Therefore, the percentage of error models is likely to be higher than 5.6% in the SAP Reference Model. The columns "Eav.", "Fav.", "Cav.", "Aav." and "Cycle" give descriptive statistics for the EPCs of a particular branch. "Eav." is the average number of events, "Fav." is the average number of functions (i.e. business activities), "Cav." is the number of connectors (for routing), "Aav." is the average number of arcs, and "Cycle" indicates whether the model contains a cycle. These numbers show that there are great differences between the different branches.

**Table 1. Errors in Specific Branches of the SAP Reference Model**

| Branch | Model | % | EPC | % | *Eav.* | *Fav.* | *Cav.* | *Aav.* | Cycle | Error | % |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Asset Accounting | 461 | 4.7% | 43 | 7.1% | 13.9 | 4.0 | 5.2 | 23.3 | 0 | 7 | 16.3% |
| Benefits Administration | 50 | 0.5% | 6 | 1.0% | 9.5 | 3.3 | 5.8 | 19.7 | 3 | 0 | 0.0% |
| Compensation Management | 122 | 1.2% | 18 | 3.0% | 7.6 | 3.4 | 3.3 | 13.7 | 3 | 1 | 5.6% |
| Customer Service | 402 | 4.1% | 41 | 6.8% | 16.5 | 3.6 | 9.0 | 29.5 | 3 | 1 | 2.4% |
| Enterprise Controlling | 599 | 6.1% | 22 | 3.6% | 14.3 | 10.1 | 6.1 | 32.1 | 0 | 3 | 13.6% |
| Environment, Health, Safety | 102 | 1.0% | 19 | 3.1% | 3.5 | 2.7 | 1.2 | 7.0 | 0 | 0 | 0.0% |
| Financial Accounting | 614 | 6.2% | 54 | 8.9% | 13.0 | 4.0 | 5.1 | 21.8 | 0 | 3 | 5.6% |
| Position Management | 4 | 0.0% | 0 | 0.0% | 0.0 | 0.0 | 0.0 | 0.0 | 0 | 0 | n.a. |
| Inventory Management | 184 | 1.9% | 3 | 0.5% | 15.0 | 7.0 | 6.0 | 28.0 | 2 | 0 | 0.0% |
| Organizational Management | 37 | 0.4% | 5 | 0.8% | 12.0 | 3.0 | 6.6 | 24.0 | 3 | 0 | 0.0% |
| Payroll | 541 | 5.5% | 7 | 1.2% | 5.7 | 3.1 | 2.1 | 11.4 | 0 | 1 | 14.3% |
| Personnel Administration | 15 | 0.2% | 4 | 0.7% | 7.3 | 1.5 | 4.0 | 12.3 | 0 | 0 | 0.0% |
| Personnel Development | 60 | 0.6% | 10 | 1.7% | 8.7 | 2.5 | 4.4 | 15.6 | 3 | 1 | 10.0% |
| Personnel Time Management | 87 | 0.9% | 12 | 2.0% | 10.8 | 3.0 | 5.3 | 19.5 | 1 | 2 | 16.7% |
| Plant Maintenance | 399 | 4.1% | 35 | 5.8% | 20.5 | 4.2 | 11.4 | 37.8 | 9 | 1 | 2.9% |
| Procurement | 444 | 4.5% | 37 | 6.1% | 6.7 | 3.5 | 2.7 | 12.4 | 0 | 2 | 5.4% |
| Product Data Management | 366 | 3.7% | 26 | 4.3% | 4.5 | 5.4 | 2.2 | 13.7 | 0 | 0 | 0.0% |
| Production | 296 | 3.0% | 17 | 2.8% | 8.8 | 3.0 | 2.9 | 13.7 | 0 | 1 | 5.9% |
| Production Planning | 194 | 2.0% | 17 | 2.8% | 5.7 | 2.9 | 3.0 | 11.5 | 0 | 0 | 0.0% |
| Project Management | 347 | 3.5% | 36 | 6.0% | 8.5 | 3.8 | 2.2 | 14.0 | 0 | 0 | 0.0% |
| Quality Management | 209 | 2.1% | 20 | 3.3% | 20.5 | 3.8 | 11.7 | 37.8 | 1 | 1 | 5.0% |
| Real Estate Management | 169 | 1.7% | 6 | 1.0% | 12.7 | 6.5 | 7.3 | 27.0 | 1 | 1 | 16.7% |
| Recruitment | 56 | 0.6% | 9 | 1.5% | 7.4 | 2.6 | 4.1 | 13.8 | 3 | 0 | 0.0% |
| Retail | 842 | 8.6% | 1 | 0.2% | 7.0 | 5.0 | 2.0 | 11.0 | 0 | 0 | 0.0% |
| Revenue & Cost Controlling | 568 | 5.8% | 19 | 3.1% | 16.5 | 10.2 | 7.9 | 36.0 | 1 | 1 | 5.3% |
| Sales & Distribution | 703 | 7.1% | 76 | 12.6% | 10.6 | 3.1 | 4.3 | 16.6 | 0 | 1 | 1.3% |
| Training & Event Management | 95 | 1.0% | 12 | 2.0% | 13.0 | 2.7 | 6.2 | 22.2 | 0 | 1 | 8.3% |
| Travel Management | 116 | 1.2% | 1 | 0.2% | 24.0 | 7.0 | 16.0 | 48.0 | 0 | 0 | 0.0% |
| Treasury | 1761 | 17.9% | 48 | 7.9% | 10.5 | 3.5 | 4.5 | 18.1 | 0 | 6 | 12.5% |
| All 29 Branches | 9844 | 100% | 604 | 100% | 11.5 | 4.0 | 5.2 | 20.8 | 33 | 34 | 5.6% |

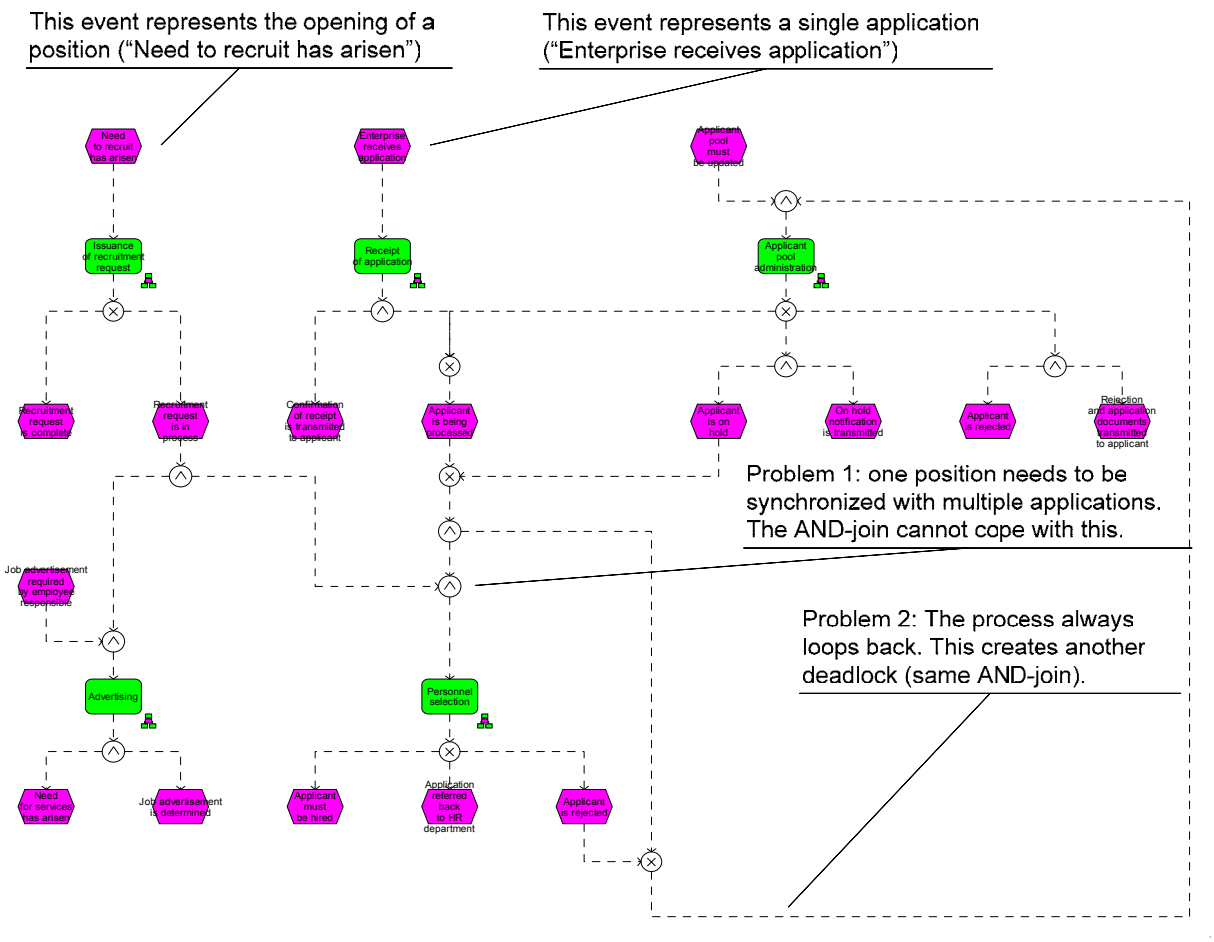


**Figure 2. Problems in the EPC of the Recruitment Branch of SAP Reference Model.**

Some of the errors are result of too little care in the design process or misunderstandings. The EPC in Figure 2 is taken from the Recruitment branch of the SAP Reference Model. It defines the announcement of a vacant position and the handling of applications. The process starts with the decision to announce a vacant position (event left at the top) and individual applications (event at the top in the middle). As these two parts of the model later join with an AND, it is implicitly assumed that always precisely one application is received, i.e., the second application will deadlock. Furthermore, there is a deadlock if there is an applicant is put in the waiting state. These errors cannot be detected with our automatic analysis approach. A tool such as ProM (www.processmining.org) is able to discover these errors since the combination of initial events can be defined. This example illustrates that the 5.6% really has to be considered as a lower bound for the number of errors in the SAP Reference Model.

The errors that we found point to a potential for improvement of the quality of the SAP Reference Model. In particular, it is critical to use models without errors for the training of staff. More serious even, it is risky to base software buying decisions or customization decisions upon erroneous EPCs. As such the errors reveal a mismatch between the software and the best practice recommendations in the reference model.

Some people may argue that in the current setting the Reference model does not directly drive the SAP system and therefore the errors are less relevant. However, such an argument reveals another weakness of the Reference model. Although it was marketed as an essential part of the SAP system, it remained rather disconnected. If one compares SAP to a VCR (i.e., a videocassette recorder), then

the Reference model was intended to be both the functional specification and the manual of the VCR. However, its role as a functional specification is problematic because of the disconnection between the system and the EPCs. Moreover, its role as a user/administrator manual is also problematic because of the many inconsistencies. Note that SAP is much more complicated than a VCR. Therefore, the lack of a good reference model is a serious problem that needs to be addressed urgently.

## When And Why Do Designers Make Errors?

Given these results we went a step further to identify what determinants have an impact on the error probability. It is of special importance to find the reasons for errors since an increasing number of information systems are model driven, e.g. workflow systems are driven by executable process models. Errors in models are immediately forwarded to the implementation and lead to faulty systems. In the context of our analysis we investigated 15 characteristics of the models and their impact on errors. We found that the size and the complexity of the models significantly increase the error probability. This fact can also be recognized in Table 1 as the Treasury branch with relatively large models shows more errors than e.g. the Environment, Health, and Safety branch with relatively small models. Based on just a few characteristics it is possible to correctly predict whether an EPC will have an error of not in 95% of the cases. These findings are valuable for the support of professionals in the design phase. This calls also for respective support in business process management tools in order to increase the quality of models.

Further details on the analysis method and the errors that we found, refer to:

J. Mendling, M. Moser, G. Neumann, H.M.W. Verbeek, B.F. van Dongen, and W.M.P. van der Aalst. A Quantitative Analysis of Faulty EPCs in the SAP Reference Model. BPM Center Report BPM-06-08, BPMcenter.org, 2006. This report can be obtained via BPMcenter.org. The used tools are open source and can be downloaded from www.processmining.org (ProM) en www.yawl-system.com (WofYAWL).