DISCUSSION

Authors: Bamkin, Macrae

Carleton Sperati: Is the documentation based at all on SGML?

Roger Bamkin (author's response): Yes, but not completely.

Philip Sargent: How many finite element programmes do you have interrogating your database in the middle of the night? Are they all located in Derby?

Roger Bamkin (author's response): The database described here is a design for a replacement to our existing system. The existing database has of the order of thirty programs which access the data from any of the Rolls-Royce sites in the UK or USA. The majority are finite element analysis proograms and require access to data approximately a thousand times a week. Some of these programs may be *accessing* particular interpolations of the data several million times within the same program run. Those programs that are not interactive (usually because they are heavy users) are likely to run in the early hours of a morning, and there is therefore no opportunity to ask for human expertise. The calculations either get the data they require or they fail (with a large cost in wasted preparation and computing time).

Norman Swindells: In the STEP model for materials the starting point is the material product, but in your system you have taken the starting point for classification to be chemical composition, a view point we rejected in the STEP model. Could you comment on the different views?

Roger Bamkin (author's response) - If a database allows but one method of defining a material, it is appropriate to choose the process used to create the material as the basis of that method. In our paper we allude, under the heading "Data Usage", to the existence of databases other than COMMIT, the principal of which is our processes database. As in the STEP model, the relationship to processes defines the **simplest** material. A disadvantage of this approach is that higher level concepts like "the aluminium alloys" cannot be defined. The introduction of hierarchies provides the functionality of being able to store information relating to higher concept levels, for which properties are generic, as well as offering a route for finding information. The hierarchies used in the presentation of tabular and other information in conventional publications are normally based on chemical composition, in line with our belief that composition is the most important hierarchy. However, this view is much debated. I have argued that the concept of hierarchy could be built into STEP, but this would increase the cost to each implementor and is certainly outside the short term scope of STEP. In communications received since this paper was written, Demaid argues that data exchange should use the simplest type of data structure. A decision on this will be needed in the next phase of STEP.

Hermann Kröckel: The result of a highly specified query, such as that shown in the third

example of Table 1 of your paper, is usually that the database does not have a data point exactly meeting that query. Problems of this type can be solved by designing the database to associate data with knowledge of mechanisms or their mathematical representation, i.e. models. This concept (for instance applied in the Petten High Temperature Materials Database, HTM-DB,) enables the computation of values not found, by interpolation and limited extrapolation from available data on the basis of mechanistic representation.

Roger Bamkin (author's response): Your comment pertains to the evaluation process which is one use of a database, but of course, is not the database or its management system. In order to answer your question effectively I need to broaden the scope in order to place my reply in context. The database described excludes (at present) mechanical test data. Validated test data are evaluated using statistics, experience and mathematical models. The resulting design data (or equations) will reside in this database and any interpolation which is theoretically possible will be permitted by the property structure. As the data should have been extrapolated within the evaluation process, when all the metadata were available, it would be dangerous to extend the extrapolation further. It is possible to argue that there may be as yet unconsidered relationships between different properties of single materials and between the same properties of related materials that would enable further extrapolation. However, this would require an evaluation process which would be done by exception, as there is insufficient time between the data being requested and the data being required (fractions of a micro second) to perform these calculations "on the fly".

The third example referred to was intended to illustrate an enquiry that would normally be satisfied - by interpolation (on numeric variables) and exact matching (character or integer variables). However, having maximised all possibilities for extrapolation and modelling there will still be cases where no data exists. Generally, where all variables are present except say "environment=marine" then the default environment would be used ("environment=air").