Poster Paper in Proceedings of the 16th Annual NACCQ, Palmerston North, New Zealand July, 2003 (eds) Mann, S. and Williamson, A. www.naccq.ac.nz

Data modelling approaches: divergence in practice

dr Samuel Mann
School of Information Technology and Electrotechnology
Otago Polytechnic
asewell@tekotago.ac.nz

In 2003, the authors were responsible for much of the third year database course in Otago Polytechnic's Bachelor of Information Technology. Mid-semester course evaluations indicated that students were interested in the different approaches to data modelling practice described by the authors. In response to this we established a small experiment in class over two weeks. The two streams of the course would separately undertake workshops on data modelling led by each of the two authors on the same described problems. Two problems were described: a staff appraisal system, and a quote generation system (shown here SM topright, AS lower).

Andrew's method aims to develop a relational database schema. Initially he searches the problem statement for entities, attributes and relationships - they appear as nouns with instance sets, descriptors or qualifiers and relationships between noun instances, respectively. This stage should follow brainstorming and no possibilities discounted. Secondly he reduces the entity list by removing irrelevancies and merging synonyms - the relationship list alters by following the entity modifications. Only applicable attributes are allowed to remain. Lastly, he builds a relationship schema from remaining individual relationships, resolves many-to-many relationships and circular references - additional relationships and entities may need to be discovered. He always tests his design with realistic data before finalisation.

Andrew Sewell

With a background in landscape ecology, Sam's practice can be described as holistic. The class read though the problem description and talked about the operation of the business. To aid discussion a diagram of the system was drawn, cases were then developed to test the system eg "quote for a door with 2 panes of glass" and questions such as why might we want make something in house even though it might be more expensive. The first ERD was compiled from the system diagram, and then iteratively refined by driving the cases though it, most modifications being towards the two sided relationship approach described in Hay's data model patterns.

As might be expected, the ERD's differ. Andrew's is more complete in terms of attributes etc (SM "I would have gotten there eventually") and Sam's is more complete in terms of nuances of the system (AS "I would have picked them up as I started to develop it"). A good experience, and one worth repeating.