

Expert judgement in cost estimating: Modelling the reasoning process

Christopher Rush & Rajkumar Roy

Department of Enterprise Integration
School of Industrial Manufacturing and Science
Cranfield University
Cranfield, Bedford,
MK43 0AL,
United Kingdom.

Tel: 01234 754072

Fax: 01234 750852

Email: c.rush@cranfield.ac.uk and r.roy@cranfield.ac.uk

Abstract

Expert Judgement (EJ) is used extensively during the generation of cost estimates. Cost estimators have to make numerous assumptions and judgements about what they think a new product will cost. However, the use of EJ is often frowned upon, not well accepted or understood by non-cost estimators within a concurrent engineering environment. Computerised cost models, in many ways, have reduced the need for EJ but by no means have they, or can they, replace it. The cost estimates produced from both algorithmic and non-algorithmic cost models can be widely inaccurate; and, as the work of this paper highlights, require extensive use of judgement in order to produce a meaningful result. Very little research tackles the issues of capturing and integrating EJ and rationale into the cost estimating process.

Therefore, this paper presents a case with respect to the wide use of EJ within cost estimating. EJ is examined in terms of what thought processes are used when a judgement is made. This paper highlights that most judgements are based on the results of referring to historical costs data, and then adjusting up or down accordingly in order to predict the cost of a new project. This is often referred to as analogy. The reasoning processes of EJ are identified and an inference structure has been developed, which represents an abstraction of the reasoning steps used by an expert as they generate an estimate. This model has been validated through both literature and interviews with cost estimating experts across various industry sectors. Furthermore, the key inferences of the experts are identified. These inferences are considered as those where many of the assumptions and expert judgements are made. The thesis of this paper is that through modelling the reasoning processes of EJ, it becomes possible to capture, structure, and integrate EJ and rationale into the cost estimating process as estimates are being generated. Consequently, the

rationale capture will both improve the understanding of estimates throughout a product life cycle, and improve management decisions based upon these cost estimates.

Keywords: Cost Estimating Rationale, Expert Judgement, Analogy based cost estimating, Cost Estimating Knowledge, Inference Modelling.

1. Introduction

In this article, the use of Expert Judgement (EJ) is examined with respect to generating cost estimates. The approach taken is one of understanding the thinking and reasoning processes that experts use as they refer to historical data to make judgements. Understanding the reasons and logic behind estimates is not a trivial task. This is because cost-estimating knowledge is complex and the sources are varied, as depicted in Figure 1 below.

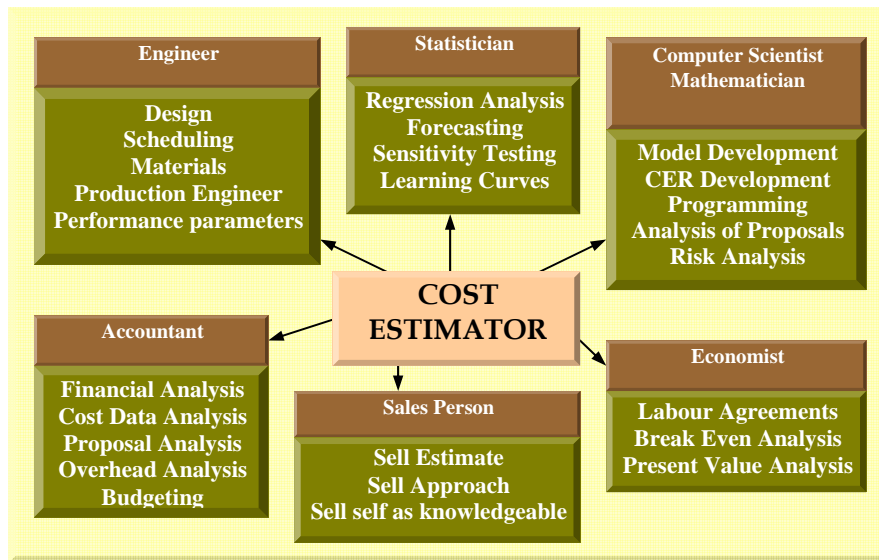


Figure 1: Skills and knowledge of cost estimating [1]

Cost estimators constantly apply a combination of logic, common sense, skill, experience, and judgement, in order to generate a final estimate that is timely, relevant, and meaningful. It is difficult for non-cost experts to understand how a final estimate is derived, and visualise the underlying reasons behind the many assumptions used throughout the process. Therefore, in order to make tacit knowledge (hidden) more explicit and useable, the knowledge needs to be

captured as the estimate is generated. This will facilitate the understanding of cost information within a CE environment and assist concurrent cost estimating.

The main concerns of this research are recognising the widespread use of EJ, and attempting to integrate the rationale behind judgements into the cost estimating and modelling process. The potential benefits for capturing cost estimating rationale are manifold, some of which are outlined below:

- Explicitly represented rationale can help individual estimators clarify their thinking about the generation of an estimate;
- The use of EJ becomes a more structured process;
- The 'hidden knowledge' of experts is captured for future reuse by non-experts and experts;
- The reasoning behind the decisions becomes available for others to critique and learn from, and;
- Existing estimates that addressed similar requirements can be retrieved, understood more easily and modified to meet current estimating needs.

In order to realise these benefits; this paper presents the development of an inference model that describes the reasoning steps used by an estimator as they make comparisons or analogies. An inference model differs from that of a process model. A process model defines steps in a process where as an inference model describes the reasoning steps of an expert carrying out a process. By understanding the reasoning patterns of cost estimators, it becomes possible to develop a software Cost Estimating Rationale Capture (CERC) model that can guide, structure, and capture

their rationale during the process of generating an estimate. This CERC model can then be integrated with other cost estimating software tools.

The remainder of this paper discusses how the inference model was developed. In section two, related research is reviewed, with particular reference to the uses of cost models, expert judgement, and analogy based cost estimating. Section three presents related work of the authors, which provides context, background information and knowledge to support the research described within this paper. Section four introduces the inference model, illustrates it, and describes how it was validated through questionnaires and interviews with experts from industry. Section five discusses related issues, limitations and the direction of future research, before concluding in section six. This paper assumes prior knowledge of cost estimating techniques and methods see Rush and Roy [2] for an overview.

2. Related research

The following literature survey discusses the use of EJ within the cost estimating process, with particular reference to EJ and commercial cost models. It also discusses how drawing comparisons or analogies are recognised as the main method used by experts as they make predictions. Finally, the current approaches to formalising analogical approaches are discussed.

2.1 Expert Judgement

Subjectivity is an issue that surrounds the compilation of all cost estimates and the use of EJ is unavoidable whether complex cost models are used or simple spreadsheets [3, 4, 5, 6]. By nature,

an estimate is a prediction of what experts think something should cost. EJ, although not a cost estimating technique, is widely used and acknowledged as necessary for generating estimates [6, 7, 8]. To be successful, the expert needs to have many years of experience. This method is obviously prone to bias; the limitations can be summarised as:

- Subjective;
- Risky and prone to error;
- Three experts with the same starting information will provide different cost estimates;
- Use of expert judgement is not consistent and an unstructured process;
- Prone to bias: personal experience, political aims, resources, time pressure, memory recall;
- The reasoning is known only to the owner of the estimate;
- Estimate reuse and modification is difficult;
- Difficult to negotiate effectively with customers;
- Difficult to quantify and validate the estimates;
- Estimate depends on level of experience;
- Experts leave the company – knowledge loss;
- Difficult to provide an audit trail, and;
- Estimates are black box in nature;

However, there are advantages to using EJ, such as:

- Quick to produce;
- Requires little resource in terms of time and cost, and;
- Can be as accurate as other more expensive methods [7, 9].

Perhaps the most formal and rigorous method for capturing EJ is the Delphi technique [10, 11]. This method attempts to capture expert opinion through a group of experts. The major drawbacks are related to its practicality. The first is related to the time needed to obtain the group opinion, and the second is related to the number of experts required to produce worthwhile results. Estimators make many qualitative judgements as they generate estimates and are often under time constraints and working with limited amounts of information. Furthermore, the Delphi technique does not attempt to capture the reasoning process of how an expert made their judgement, which is a main aim of this research.

2.2. Commercial cost models and EJ

Examples of widely employed cost models are: parametric or algorithmic models, and non-algorithmic models, such as Fuzzy Logic (FL) and Neural Networks (NN), and Case Based Reasoning (CBR). These tools are mostly computerised, which contain algorithms, rules, inferences and mappings. Once calibrated to a particular environment, it can be argued that cost estimating becomes more scientific because the process is more repeatable. Nonetheless, these models can only aid the cost estimator and cannot replace them [5]. Cost estimators have to use their judgement concerning the validity of calibration data, and during the input of parameter values. For example, parametric cost models are often built on underlying assumptions and relationships between variables, which do not necessarily reflect reality [4, 5, 12]. Thus, it is the cost estimator and their expertise that ultimately controls the output of any cost model. Furthermore, the judgements and assumptions used can influence the results significantly. Through the author's own research, concerning a commonly used cost model called PRICE H, the use of expert judgement during the calibration process was extensively used [3, 13]. Thus,

although cost models are designed to facilitate the generation of estimates, they can only reflect a realistic cost with the input of expertise and judgement of the users.

It should be noted that companies feel more comfortable with the use of algorithmic and computerised models [14, 15, 16, 17], than they do with EJ (Hughes [7]). Despite the fact of enormous error ranges. For example, within the software cost estimating community errors of between 85 – 770%, [7, 8, 9], are reported. Pine [18] reports on estimating errors of hardware projects, although not so large, significant nonetheless. Hughes [7] argues that companies should not take such a negative view with respect to the use of EJ. Rather, that companies should acknowledge it, and develop cost models and information systems that attempt to support it. Hughes' work does not attempt to do this.

Thus, very little research discusses how an expert uses judgement or how EJ can be better integrated into the cost estimating process. None of the research attempts to make this process explicit or attempts to capture cost estimating rationale as an estimate is generated. The reasoning and rationale behind the input values into cost models is more often than not, not recorded and hence is lost to other experts and non-experts alike.

2.3. Analogy and expert judgement

The idea of using analogies as the basis for estimating is not new. Boehm [19] suggested the use of analogies as a means for estimating some 20 years ago. It is widely accepted that the most common way in which experts produce cost estimates and make judgements is through the use of analogies or comparisons [7, 8, 20, 21, 22, 23, 24].

Within the field of cognitive science, general models of analogical reasoning are provided [23, 25]. In broad terms, they describe how a reference product (source) of which the details are known, is first identified for use as a comparison to the current product (target) of which the details are not fully known. The similarities or differences are then matched, and then transferred from the source to the target. Finally, a justification is made to explain why the source product was used. This type of process has been used by the US air force for over twenty years for predicting the repair rates of aircraft parts [24]. NASA recognises a similar model and provides guidelines for using analogy as a means to estimate on their web site [26]. These models provide an abstraction of the process but do not describe the types of inferences and thinking processes used by the experts. Furthermore, these models do not describe the levels of abstraction that experts use as they identify whether a case is suitable for use as a comparison or not [27].

Almost all research with respect to the use of analogy as a basis of estimating, revolves around the creation of analogy based tools i.e. Case-Based Reasoning (CBR) [8, 20, 22, 28, 29, 30]. The attraction being that CBR more naturally resembles the thought processes of experts. A CBR system stores and organises past situations, then chooses those similar to the problem at hand and adapts a solution.

Bashir and Thompson 2001 [21] provide an alternative method of comparison using a more manual approach. In this approach, they suggest the use of a pairwise comparison table and an eigenvector approach. These enable the user to define a similarity measure relating to the productivity of historical products to predict the design effort of new projects. Their work

doesn't explain the reasoning processes used by estimators but does offer a method of assessing how projects are being compared.

Formalising and computerising the process of analogical reasoning into a CBR system is a complex process. There are many issues concerned with the retrieval, matching and measurement of similarity between past and current cases. Furthermore, a CBR system requires a number of past cases in order to be effective. This is a problem when a company does not have such a number of past cases that can be used. For example, within the military aircraft industry, this is often the case. Much of their historical data spans over 50 years and retrieving this is not a trivial task. In addition, each new generation of military aircraft is often a step up in technology from its predecessor. Regression based tools are good at predicting within the bounds of similar projects as are CBR tools. If a new technology is introduced, the systems fall down or are severely limited in their predictive capability. In a CBR tool, a similar match is not found, and extrapolating beyond the range of data points in a regression-based model leads to a spurious estimate indeed. Thus, the use of EJ is unavoidable.

Therefore, the authors' approach to 'analogy' based cost estimating, is first to understand and capture the reasoning processes used when experts make comparisons to generate an estimate. By understanding these issues, it becomes possible to model the expert behaviour, guide the process of using EJ, and capture rationale where necessary. As mentioned previously the benefits of this are manifold. Before the inference model of the thought processes is presented, the related work of the authors is introduced to provide context, background information and knowledge, to support the current research described later in this paper.

3. Challenges in cost estimating knowledge capture

The research within this paper was initiated by Cranfield University and a European Military Airsystem manufacturer. It is part of a three-year research project, in which the sponsoring company's cost modelling activities of new military aircraft is being examined. The main aim of the project was to examine both the quantitative and qualitative issues of cost modelling with the objective of integrating the two types of knowledge [3, 13].

- Quantitative knowledge is defined as the elements of known cost and product structures, which form the basis of a cost estimate and are measurable. For example, the product breakdown structure (PBS) of an air system can be defined and decomposed into a hierarchical breakdown of both its major and sub structures and systems. The masses of parts and material types are other examples of quantitative knowledge.
- Qualitative knowledge is defined as the assumptions and judgements that cost estimators and engineers make during the generation of an estimate. These assumptions and judgements are related to how an estimator refers to past projects to serve as the basis for generating a new estimate.

New military air systems can take as long as, and more than, twenty years to develop from the concept evaluation of customer requirements through to the final production [31]. This time span makes it extremely difficult to predict the cost and risk of a multi-million pound product during the concept evaluation phases. Furthermore, many authors agree that 70-80% of a product cost is committed during the concept phase [14, 15, 16]. Making a wrong decision at this stage is extremely costly further down the development process.

Nonetheless, the sponsoring company's cost estimators are required to predict the full life-cycle cost for such options. They need to make assumptions about future cost savings resulting from: new manufacturing processes, the use of more advanced technologies, new materials, expected improvements in design and development processes, and changes within and outside the business environment. To achieve this, they often make comparisons and references to historical projects in order to predict the cost of the new project. Furthermore, the cost estimators need to communicate and work constantly with their IPT (Integrated Project Team) members in order to help produce estimates that are more meaningful.

The authors discovered that as the cost models are developed, much of the reasoning and logic is understood only by the experts. Where project time scales are long, the reasoning and logic behind the estimates is often lost, and it is difficult for both non-cost experts and other experts to understand how a final estimate has been derived. The underlying reasons behind the many assumptions used throughout the process are not visible. This fact influenced the decision to capture this knowledge and rationale in order to make it more explicit; thereby, making the final estimate more meaningful and reusable.

Capturing this rationale and identifying the knowledge intensive processes is not a trivial task. It is difficult for experts to explain what it is that they know [32, 33, 34]. In previous research, the authors investigated the knowledge capture of experts related to the cost modelling of a new product [3, 13]. In this work, expert knowledge was captured and the knowledge intensive areas of a cost modelling process were identified. The cost modelling process was mapped using an

IDEF₀ approach, and the knowledge intensive areas were identified and captured through the use of a novel technique known as Knowledge = Expert – Novice (KEN) [35]. Within this approach, the Novice (author) used the data and tools of the cost-estimating experts to produce an estimate. As difficulties arose, the Novice called on the expert for the solution. As the Novice works through the required activities and tasks the knowledge requirements are documented [3, 13]. Figure 2 below illustrates the knowledge identification phase of this research.

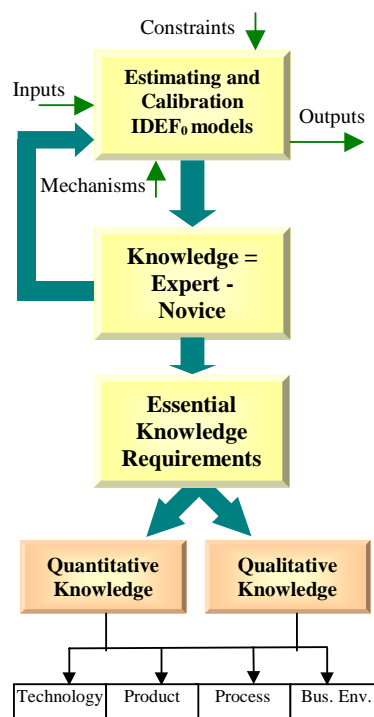


Figure 2: The knowledge identification methodology [3, 13]

Because of this research the knowledge intensive areas were identified, and a model depicting the current AS-IS process was developed, which facilitated the development of the future, TO-BE cost estimating process model (see Figure 3).

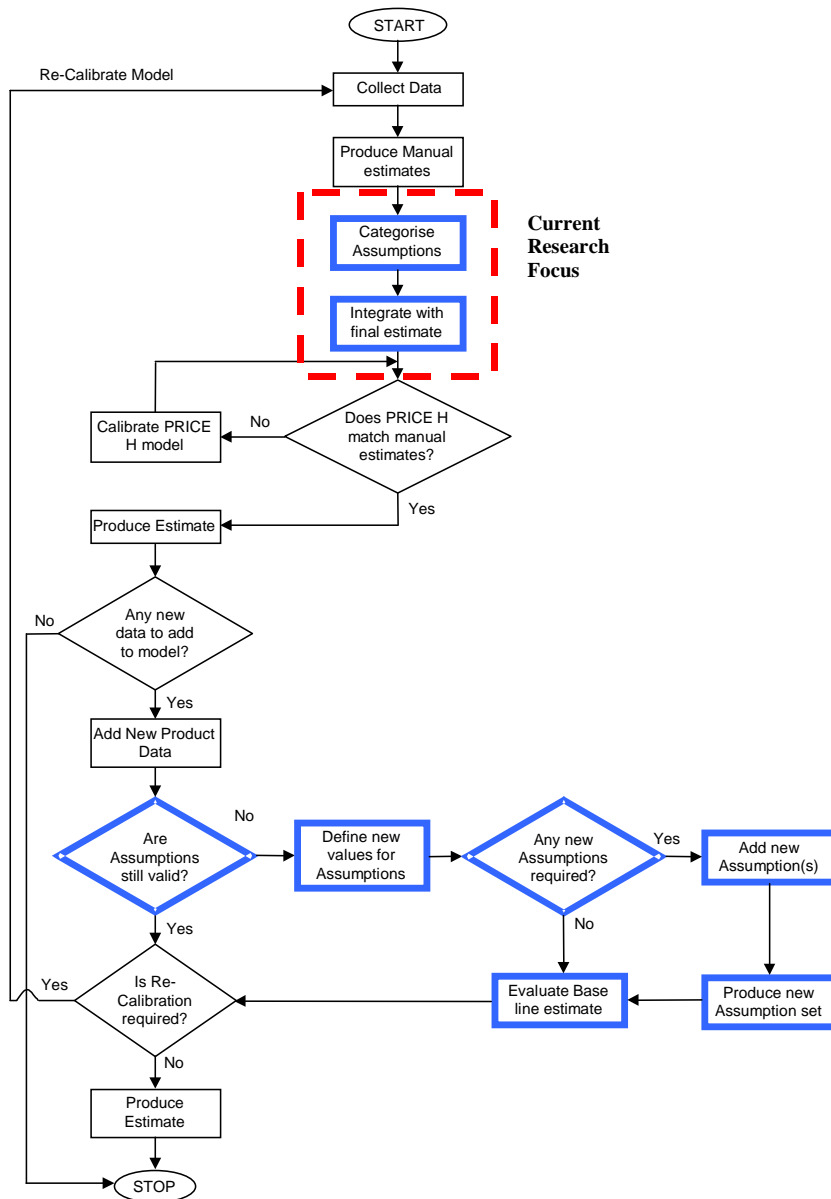


Figure 3: TO-BE calibration process and research focus (3)

Figure 3 illustrates where changes to the existing cost model process are being introduced by the bold diamonds and boxes. The information described within this paper is depicted within the area of the dashed-bordered box. As mentioned within the introduction, the objective within the research focus area was to understand the reasoning and thought processes of the estimators as

they make assumptions, judgements and decisions. The remainder of this paper describes the inference model in more detail; also, how it was validated through semi-structured interviews with cost estimating experts.

4. Developing the Inference model

4.1. Initial model development

After the knowledge intensive areas of the cost modelling process had been identified, the authors set about capturing the reasoning processes used by the cost estimators. An initial inference model was developed using the CommonKADS notation [34]. It should be noted that the inference structure has not been developed to drive a knowledge-based system. Rather, to provide a basis for developing a software model that mimics and guides the expert thinking process. The initial inference structure was developed based on the authors' understanding of the domain in which the knowledge identification was carried out and through existing literature.

4.2. Inference structure validation

In order to validate the model a questionnaire was designed to gather further information and model the expert reasoning process more closely. In order to provide some structure to the questionnaire the respondents were required to refer to an estimate they themselves generated. The inference structure in Figure 5 was validated through interviews with 11 experts from aerospace, automotive and defense industries and estimating backgrounds. Their years of experience ranged from 5 to 32. The projects ranged from cost estimating new airsystem concepts to the disposal of nuclear waste plants. The questionnaire is illustrated in Table 1.

Table 1: Semi structured interview questionnaire

Section 1: The Context	
1.	What type of product was the estimate for (mechanical, systems, etc.)
2.	What was the stage of product development? (Development, Manufacturing)
3.	What estimating techniques were you using? (e.g. parametric, detailed analysis)
4.	When was the estimate made?
5.	What was the purpose of the estimate? (e.g. ROM, Budget, detailed, fixed price etc.)
6.	How would you rate the difficulty of this task? (On a scale of 1 to 5, 1 being very easy and 5 being very difficult)
7.	How many years of cost estimating experience with respect to generating estimates had you (at the time of estimate)
8.	How often were you producing estimates? (at the time of estimate)
Section 2: The process	
1.	Where does the work request come from?
2.	What are the basic data you require, and the main things you look for? And why?
3.	What sort of constraints do you often face (example: time, etc. etc.)?
4.	At what point could you say that you understand the new product and the cost estimating requirements? Please give an example from your case study.
5.	What tells you that you have all the requirements you need, whom would you talk to for any advice? How often do you consult others?
6.	What do you do to identify and prioritise requirements?
7.	What do you have to do to organise the new product data for preparing the estimate?
8.	At what point do you begin searching for source or reference data?
9.	What difficulties do you encounter when searching for source data?
10.	What do you look for in the source data? Please give examples from your chosen case study.
11.	What tells you that you have identified a suitable candidate to use as a comparison?
12.	What do you do to when comparing the source data and the new product data? Can you provide examples? (what criteria do you use? And why?)
13.	What assumptions do you make when comparing the products? Could you please give examples?
14.	How would you use the costs from the source data?
15.	What do you do when you make adjustments to account for the differences between the source data and the new product data?
16.	What sort of assumptions do you use when adjusting the costs? Could you please give examples?
17.	Do you record and explain the rationale for the adjustments made? If yes, can you say how?
18.	How often would you review any assumptions made?
19.	Do you understand the cost impact of a change in any of the assumptions used?
20.	What are the potential areas where your estimate may differ from another person's estimate?
21.	Could you please explain the thinking process you go through when making a comparison, as you understand it.
Section 3: Inference model analysis	
1.	In your opinion, how do the inferences match your reasoning process? Is there anything you would change?
2.	What inferences are most assumption intensive?

Section 1 of the questionnaire was designed to gather information concerning the case study and the context in which the estimate was made. Section 2 was designed to assess what the expert did as they carried out their tasks. The questions in both Sections 1 and 2 were asked without showing the initial inference model to the interviewees. After the second set of questions had been answered, the inference model was presented. Section 3 of the questionnaire was designed to understand how the author's inference model differed from the expert's view of their reasoning process. Any discrepancies were highlighted and then later assessed through the respondents answers to the questions posed in section 2 of the questionnaire. All the interviews

were recorded and later transcribed in order to assess that their answers matched the inference model.

4.2.1. Paper Based Validation Scenario

The scenario below describes only one of the 11 cases due to paper length limitations, which illustrates how the inferences were matched with the transcribed data.

Scenario: Two cost estimators have just received a request from their line manager asking them to estimate what the cost of an aircraft wing should be costing the company. The company has been building the wings for many years. There were recorded costs against the wings but the company wanted to know how much should it be costing and what target they should be aiming to. The information received was a verbal instruction. The time allotted for the job was one month.

Table 2: Paper based validation of interviews

Domain	Explanation	Inference Model
Where does the work request come from?	We received a verbal request from the manager. The formal route is through the contracts department from an external customer and to us from project management. In this case, the managers decided that they wanted the study conducted.	RECEIVE: From Manager
What are the basic data you require, and the main things you look for? And why?	We needed to collect more data such as process plans, product details, drawings, and through visits to the shop manufacturing plants. This data helps us understand what we are cost estimating for.	OBTAIN: More Data
What sort of constraints do you often face?	Time and lack of information. Knowing the type of estimate affects the amount of detail one will use to generate the estimate. In this case, it was a detailed estimate and needed to have a good degree of accuracy. Knowing this helps to identify and prioritise the areas of a product or estimate to focus on. In this case, the estimator had one month which focuses the attention to the high cost drivers	CONSTRAINTS: Time and Information
At what point could you say that you understand the new product and the cost estimating requirements? Please give an example from your case study.	After we've seen the product, visited the manufacturing plant for a few days, looked at the line, the job, spoke to guys to find out what were the issues around producing the parts, and looked at drawings for particular parts. Also when you understand what sort of price or cost your expected to put together, what sort of economic conditions exist, and what sort of tolerances there are and how long you've got to do it.	SYNTHESISE: Understanding
What tells you that you have all the requirements you need, whom would you talk to for any advice? How often do you consult others?	Clarifying requirements through the more experienced staff and the managers right up to the senior managers. In many cases, we would also attempt to clarify requirements with the customers. For advice, you would talk to the Estimating managers, commercial managers, and project managers, as and when you needed to.	SYNTHESISE: Clarifying requirements
What do you do to identify and prioritise requirements?	If time's more important than accuracy you produce as much as you can within the time scale.	SYNTHESISE: Prioritise
What do you have to do to	We drew a product tree showing the major assemblies and all the sub-parts and how	SYNTHESISE:

organise the new product data for preparing the estimate?	they related to each major assembly.	Organise
At what point do you begin searching for source or reference data?	This happens as your firming up the requirements. It happens from the very start really.	SEARCH: Data
What difficulties do you encounter when searching for source data?	Finding usable data. Interpreting data. Asking other people to find data when they are already busy.	CONSTRAINTS: Time, people, information
What do you look for in the source data? Please give examples from your chosen case study.	It was the detailed planning's, materials, masses, sizes, and processes involved, plus access to sub contract orders that had gone out, and access to the recorded costs for the previous wings.	IDENTIFY: What they need
What tells you that you have identified a suitable candidate to use as a comparison?	It's down to prior knowledge. Experience tells you that you can read certain things across. If you know the people who done one job then a similar job done by a similar set of people should be done for the same amount of time and cost. Also, talking with other people.	IDENTIFY: Usable data
What do you do to when comparing the source data and the new product data? Can you provide examples?	Similar sized parts like the ribs, spares, stringers, and the method of construction. We look at the mass, similar materials, similar sizes, manufacturing process, construction methods and then you would expect it to cost much the same. Previous estimates will also be reviewed.	COMPARE: Similarities and differences
What assumptions do you make when comparing the products? Could you please give examples?	In this case, we assumed that similar sized parts and similar material parts will cost the same to manufacture at the reference/source data site as they would at new site, in terms of man-hours and basic costs.	COMPARE: Assumptions
How would you use the costs from the source data?	A lot of the time we bridge costs from one estimate to another and then factor for the differences. We also say what data we have used to generate an estimate (JUSTIFY).	TRANSFER: Costs from source data to target product
What do you do when you make adjustments to account for the differences between the source data and the new product data?	For example, we looked at the basic assembly cost, detailed sub assembly costs learning factors, errors on the previous estimates and then we would put tolerances and factors in for different parts. You look at the different time scales, materials, masses and then identify the differences and put in an estimate for that.	ESTIMATE: Factors and tolerances
What sort of assumptions do you use when adjusting the costs? Could you please give examples?	You look at things like batch size and the learning implications, second source learning. For example, if a part was cost based on 400 down the line you have to consider this against a part that may be made with new operatives. You have to judge where on the learning curve you should place the part. In addition, the sort of efficiencies in the work shops.	ESTIMATE: Assumptions
Do you record and explain the rationale for the adjustments made? If yes, can you say how?	Yes we do, we have to justify how we came to our cost estimate, it's part of our clearance procedure. This includes the detail about where you obtained the information to do the job and a bit of a critique. A list of all the assumptions you used, a list of source/comparison data and how you handled this. There is no set format. Risks and tolerances will also be included. This is stored on file, either computers or filing cabinets.	JUSTIFY Assumptions and rationale
How often would you review any assumptions made?	Depending on the scope of the job. On a longer job, the easy answer is to say all the time. If you only have limited time to make the assumptions you'll make them and review them before the final figure. If you have a few months, you may have made earlier assumptions with the expectations that you would be able to firm them up as you go along, and as you obtain more information. But you do really review them all the time. But definitely before you finally submit the estimate. There is nothing formal. Many of these things come down to good or bad working practice.	REVIEW
Do you understand the cost impact of a change in any of the assumptions used?	Mostly yes, this is one of the reasons for making the assumptions in the first place. Assumptions are made so that you can get on with doing the job. Process improvements, or material type assumptions should be highlighted in a risk register. One of the big assumptions we made related to this case was that a lot of the work would be done outside the factory by sub-contractors at a cheaper rate. It ended up as higher rates because the jobs were more difficult than expected. So, it's important to recognise these sorts of assumptions and recognise them as a risk. It's impossible to understand everything that might happen.	REVIEW
What are the potential areas where your estimate may differ from another person's estimate?	Just about everywhere, it's mostly related to the assumptions and factors used. Different people exercise more caution in certain areas such as error rates, broken tools, learning curves etc., and estimate more cost. At a higher level differences will occur on the number of hours people will spend per week or a month, holiday allowances. In addition, factors to allow for more profits.	N/A
Could you please explain the thinking process you go through when making a comparison, as you understand it?	We assess the source data and identify its suitability for costing the new project and then match the costs we can, we then make estimates for the differences between the source data and the new product we want to cost.	N/A

Thus, through both the interviews and later analysis the inference model was developed. The final inference model depicted in Figure 5 was accepted and considered representative by all the interviewees.

4.3. Task: Generating an estimate

Before one can capture the reasoning processes of an expert, the task they perform needs to be defined. This enables one to understand how the expert is reasoning. In this case, the task is related to what an expert does when generating an estimate, based on reference to historical projects or experience. As mentioned previously this is often referred to as analogy. The task in this case, is a complex reasoning process and needs to be decomposed in order that it can be more clearly understood (see Figure 4).

In Figure 4 the main task of generating an estimate is hierarchically decomposed into smaller tasks, this in turn is divided into even smaller tasks. The tasks in the diagram describe *what* the expert does. The task methods describe *how* an estimator completes the task.

In order to understand *what* the expert does, the task method needs to be further decomposed into subtasks i.e. Prepare (1), Estimate (2), and Review (3). Each of these subtasks has a number of associated lower level subtask methods that describe *how* the cost estimator completes each subtask. For each subtask method, the corresponding inferences used by the estimators are shown. Descriptions of inferences are provided in Section 4.4.

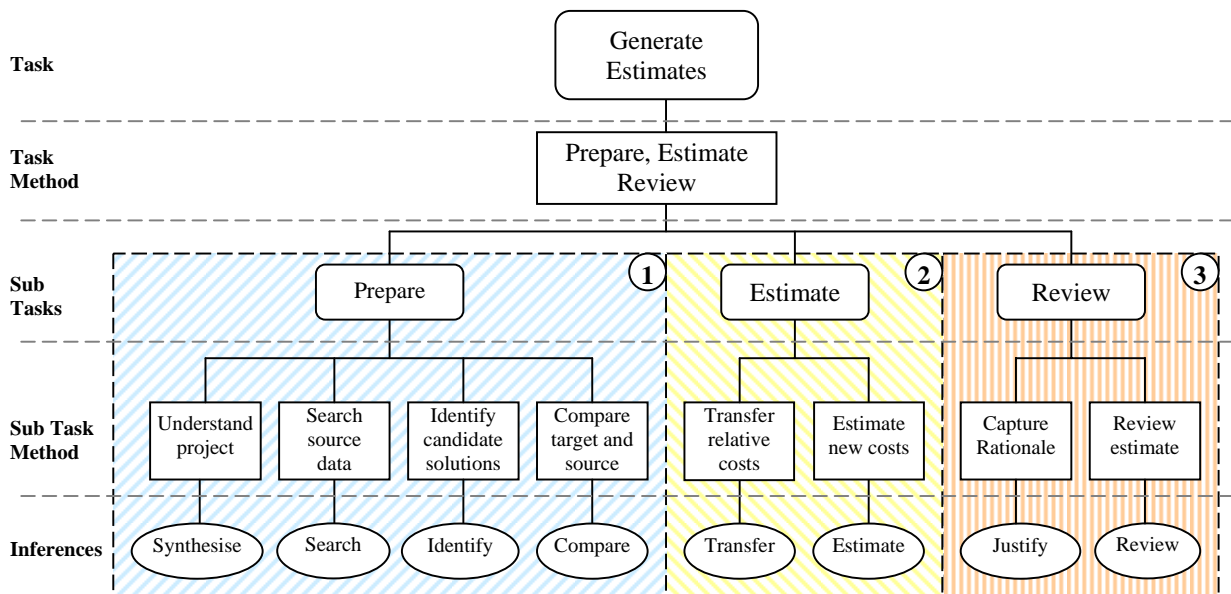


Figure 4: Task and inference decomposition for analogy based cost estimating

The main tasks prepare, estimate, and review describe the order of how the estimate is generated. For example, before estimating, the data needs to be *prepared* (1) so that the expert can *estimate* (2). Then the estimate needs to be *reviewed* (3). However, the task decomposition does not illustrate how the expert reasons during the task, it simply demonstrates the ‘what’ and ‘how’ of the cost estimator as they complete their tasks, and illustrates the inferences used. The shaded areas in Figure 4 are numbered, and correspond to the shading and numbering shown in Figure 5. This helps the reader visualise how the task of ‘generating estimates’ translates into the experts’ reasoning process.

4.4. Inference Structure

The inference structure illustrated in Figure 5, is an abstract representation of the possible reasoning steps an estimator uses as they refer to a similar product to generate an estimate. Together, these inferences form the building block of the expert reasoning process. They define the basic inference actions that the expert can perform whilst executing their tasks. The combined set of inferences represent the experts *inference structure*.

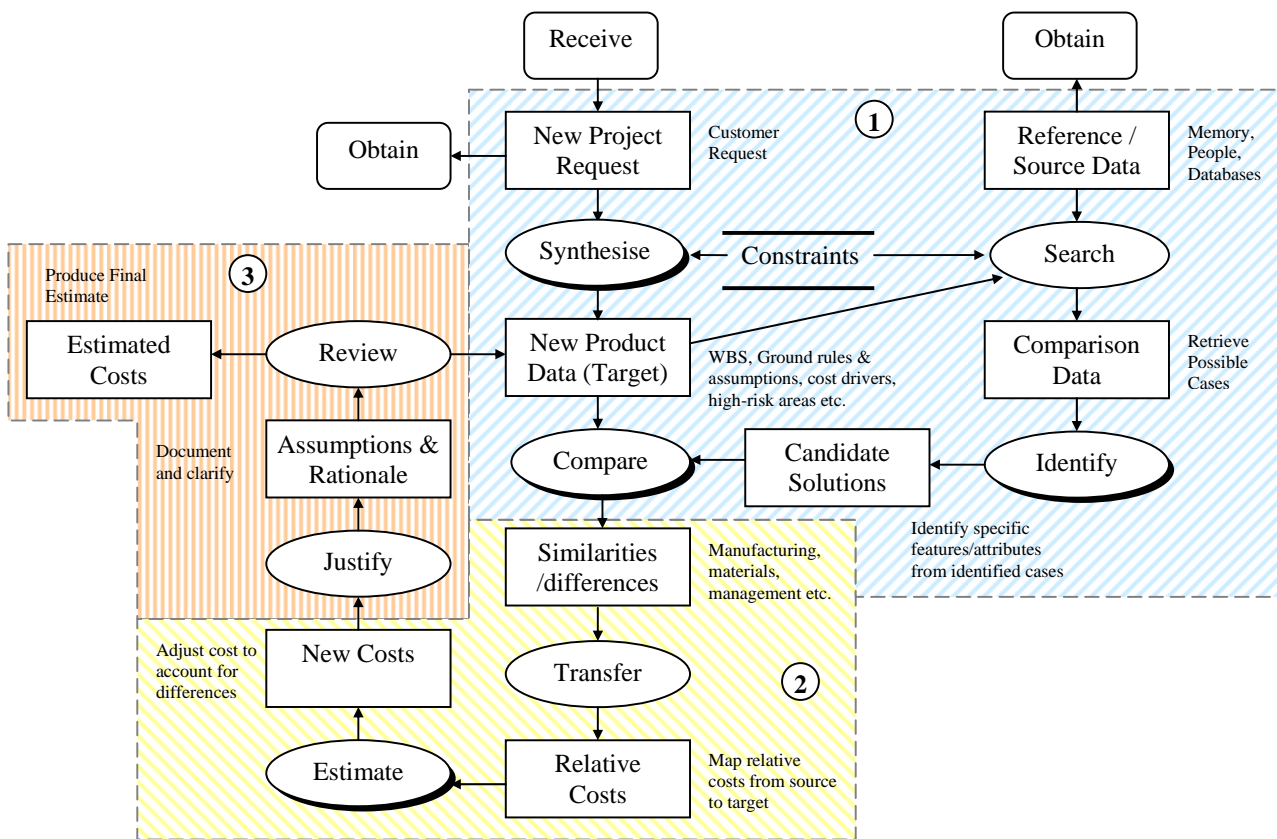


Figure 5: An inference structure for analogy based cost estimating

4.4.1. Knowledge Roles, Transfer Functions and Inferences

To model the inference structure the authors adapted the CommonKADS notation [34].

- The rectangular boxes within the model are known as Knowledge Roles (KR). The KR's describe at an abstract level, the kind of data that the estimator will infer or reason with.
- The ovals represent the Inferences (I) or the reasoning processes that the expert uses. The arrows are used to indicate input-output dependencies between the KR's and inferences.
- The rounded boxes represent Transfer Functions (TF). The TF's relate to the estimator interacting with other agents e.g. suppliers, customers, IPT members, and collaborating companies.

4.4.2. Inference structure description

The following section provides a walk through of the inference structure illustrated in Figure 5.

Prepare (1): The estimators first receive information about the new project in various forms. They *receive* a request to do the work, and data such as 3-D models, drawings and documents. From this information, the estimator deepens their understanding of what it is that needs to be cost by *synthesising* all the data and information. During the synthesis of data, the expert may need to *obtain* more data to deepen his/her understanding of the project. The estimator needs to analyse the requirements, and classify the type of product and the type of estimate they are

required to produce. They also need to understand the constraints in terms of time and resources. Once they understand what is required, they establish and document any ground rules and assumptions and identify the main cost drivers of the project. A Work Breakdown Structure (WBS) is produced and used as the framework to cost the new product.

In addition, during the synthesis of data and information the estimator normally begins *searching* for more data and *identifying* projects that can be used for comparison. For example, the search can be from memory, other people or existing databases. The estimator needs to consider what elements of the source/reference data can be used as a basis for comparison. When ready the expert will begin to *compare* the similarities and differences of source and target projects and match them. The matching can be based on different levels of abstraction with respect to features, functionality, and project management and so on. The level of detail searched for will depend on the time constraints and the type of estimate required. The searching process will happen continuously throughout the synthesis process.

Estimate (2): As the estimator compares the similarities and differences, the relative costs from the source product are *transferred* to the target product. The estimator will continuously assess the projects in order to understand what costs can be transferred and those that need to be adjusted or *estimated* for the new project. When estimating the estimator uses their experience and judgement to predict the

cost of the new product. The new costs need to be justified through cataloguing any assumptions used; however, this is not always the case.

Review (3): The final part of the expert reasoning is to *justify* and document the assumptions used. As mentioned previously, this rationale is not always captured. In addition, the estimator will continually *review* the estimate because of more information being received or obtained. Finally, the estimator may use various means to ‘sanity check’ the validity of their estimate. This can be through other people, or using other estimating techniques and tools.

4.4.3 Key inferences

The inference structure presented in Figure 5 is an abstract representation of the reasoning steps used by an estimator. In reality there are many more sub inferences used. However, the main aim of an inference structure is to get to a level of decomposition where the inferences used describe the reasoning processes to a sufficient level of detail to understand the domain. The shadowed ovals (inferences) are those identified by the experts where most of the assumptions are made during the process of generating an estimate. In future research, these inferences will be further decomposed in order to identify the knowledge intensive areas of a specific judgement. Due to the limitations of paper length, only generic descriptions of the assumption intensive inferences are presented below.

Synthesise

Operation: The *inputs* for this inference are the cost-estimating request, and the available project data. Data can include drawings, process plans, work breakdown structures. The experts obtain more data as required in order to understand the estimating requirements. The constraints are recognised, as are the high cost drivers, and high-risk areas of the project. The *output* of this inference will be the new project data prepared for comparing with the source data, such as the WBS, assumptions and ground rules and high cost drivers.

Example: Here the expert analyses, clarifies, establishes, and assimilates all the information into a format ready for comparing and producing the estimate. Typical assumptions are related to envisage process improvements, and improved communication through using CE principles.

Knowledge: The ability to identify those areas that will drive the cost, and establish the high-risk areas. This knowledge is dependent upon the project or product being estimated. In a specific domain, the high cost drivers and risk areas may have common characteristics so can therefore be captured to guide the expert within the final CERC tool.

Identify

Operation: The *inputs* are the source data and the details related to the retrieved project or product. The *outputs* are the identified features, attributes of a product, or project areas that can be used for comparison.

Example: Example assumptions would be related to the management structure, the experience of the teams, the quantity of production, the level of complexity, the functionality of the product, and the manufacturing processes used.

Knowledge: Knowledge of historical products and those areas that are commonly used as a basis for comparison would need to be identified. These would relate amongst many others to, specification, materials used, mass, type of system, manufacturing processes, assembly techniques, functionally, and productivity rates, whether VAT was used, economic

conditions, and exchange rates.

Compare

Operation: The *inputs* for this inference are the source or reference data, and the new product data (target). The *output* of this inference process will be a measure of both the similarities and differences identified by the estimator.

Example: Examples of similarity measures include: manufacturing methods of the source and target data. Assessing whether the technologies are the same, the same sorts of quantities being produced, the learning curves associated to volume, the processes required and so on.

Knowledge: This relates to the types of comparisons that are often made. For example, manufacturing methods, project management, materials, mass, technology, system type etc. A full list would need to be identified and captured and related to the domain in which the comparisons are made.

Estimate

Operation: The inputs are related to the differences identified between the source and target data. The outputs would result in an adjusted cost and new costs based on the assumptions used by the estimator.

Example: The estimator may assume that the manufacturing processes used to produce the historical product are not representative of the manufacturing processes for the new product. And may therefore assume a saving of, for example, 40%. It is here that the rationale would need to be captured in order to validate the estimate.

Knowledge: The knowledge required here would be related to the expected changes in manufacturing processes and productivity. The knowledge also includes the impact of any changes to the cost.

4.4.4 Knowledge role attributes

Each knowledge role is described in terms of the data and information that it contains. This is what the estimator infers or reasons with. This paper presents a description concerning one of the critical knowledge roles. For example, the knowledge role for ‘Comparison Data’ would contain the elements such as those described in Table 3. These are the types of data and information that an estimator will use to make a comparison. Not all of the attributes of a knowledge role will be used; the estimator will use only those that are identified as relevant to the project at hand.

Table 3: Knowledge role attributes of comparison data

1. Identify past cases with respect to each of the WBS elements
1.1. Understand and define the technical and programme details
1.1.1. Any assumptions and ground rules used
1.1.2. Unusual circumstances the product was produced under
1.1.3. Operations analysis data
1.1.3.1. Quantities
1.1.3.2. Schedules
1.1.4. Product attributes/aspects/features
1.1.4.1. Mass
1.1.4.2. Materials used
1.1.4.3. Structures
1.1.4.4. Volume
1.1.4.5. Shape
1.1.4.6. Dimensions
1.1.5. Power
1.1.6. Performance
1.1.7. Processes
1.1.7.1. Design process
1.1.7.2. System development process
1.1.7.3. Manufacturing process
1.1.8. Specifications
1.1.9. Levels of testing and analysis
1.1.10. Procedures followed
1.1.11. Constraints
1.1.12. Limitations
1.1.13. Business performance
1.1.14. Economic influences
1.1.15. Political issues
1.1.16. Sources of information
1.1.17. Validity of cost information
1.1.18. Other, please define

In summary the knowledge roles, transfer functions and inferences provide a means of representing both the reasoning processes, and the data and information that the experts reason with. This provides a base from which to develop a Cost Estimating Rationale Capture (CERC) tool to guide and capture expert judgement.

5 Discussion & Future Research

This paper began by introducing the issues related to the use of Expert Judgement (EJ) during the process of generating a cost estimate. It demonstrated that EJ is continually used throughout the process of generating estimates whether cost models are used or not. Moreover, that the use of EJ is unavoidable, because an estimate is a prediction about the future and judgements are continually required. The authors' presented the challenges of capturing cost estimating knowledge and discussed the impact EJ can have on an estimate.

The process of analogy was identified as the main method by which experts reason when generating cost estimates. Most research within the field of analogy based cost estimating focuses on the development of Case Based Reasoning (CBR) tools. The issues of integrating EJ and rationale into the cost estimating process are rarely considered. Therefore, within this paper the development of an inference structure was presented and detailed. The inference structure represents an abstract view of how an expert judgement is made when referring to historical data. The research also identifies the knowledge intensive inferences for cost estimating. In addition, the knowledge roles are defined for the domain.

The inference structure presented within this paper is one possible way of representing the reasoning processes of an expert. The structure captures the major inferences being used during the generation of cost estimates. The principle aim of an inference structure is to derive a structure that represents the domain, and that is recognised by the experts.

Experts seldom make judgements without some prior knowledge or expertise. Judgement will always be based on some reference point or prior knowledge. The judgement may not be correct but it should be possible to understand how and why it was made. Experts will often say that their estimate is based on gut feel, yet in reality this is not true [24]. What an expert considers as 'gut feel' or intuition is simply years of experience. With a few probe questions, it is possible to uncover the reasoning behind the estimate and capture the rationale.

Therefore, the main thesis of this research is that by first understanding the principles of how experts reason, it will be possible to guide, manage and capture the estimating rationale during the process of generating an estimate. This is not a wild claim as designers have been developing tools to capture design rationale for many years [36, 37, 38, 39]. Thus, the future research will focus upon the development of a Cost Estimating Rationale Capture (CERC) software tool that can be used to capture expert rationale as an estimator uses their judgement. It should be noted that the CERC tool would not be an expert system.

The CERC will facilitate concurrent cost estimating by providing a means to reuse and share rationale underlying cost estimates. The main issues that the research will consider are the

representation used, allowing estimators to express their reasoning in a natural way, while at the same time being formal enough to support useful computational paradigms. Furthermore, that the process of describing rationale should impose minimum possible overhead on the estimating process.

5.1. Research limitations

The inference structure described within this report was validated through experts from various sectors of the defence and automotive industries. Therefore, it cannot be claimed that this reasoning model is representative of generating cost estimates across all domains. In addition, the number of inferences used in the model and the level of representation has been derived qualitatively through interviews. Nonetheless, the model does depict the domain and the reasoning process to a level of detail that the cost estimators recognised.

This inference model represents the abstract reasoning processes of an expert using a reference product on which to base an estimate. However, the model does not address the issue of an expert not finding a reference product. The future CERC tool will need to consider that a reference product may not always be available and an estimate is much more subjectively derived. Nonetheless, this inference model does allow one to choose memory as the source of reference data.

Furthermore, the abstraction level of making analogies and comparisons is not addressed in the model. For example, if an expert cannot find a comparative part at a specific level of detail they

will begin to search at a higher level of abstraction to find a match [27]. These issues will also be considered in the future.

6 Conclusions

This paper demonstrates that the use of expert judgement within the cost estimating process can be structured. Not much research is being done in this area. The authors have shown that the reasoning process of experts, when generating cost estimates using analogy is generic. This is counter to the popular belief of the 'black art' or 'black box' nature of the cost estimating process. The inference structure presented represents the same reasoning process for cost estimating at different levels of product definition and detail. The same reasoning process is used whether in the conceptual stages of design or whether in the disposal life cycle phase of a project. The process or method by which an expert uses judgement does not change. Only the level of abstraction changes. In the conceptual stages of project development, the comparisons are made at a much higher level of project definition, and as the project moves into the later stages of development, the comparisons become increasingly detailed.

The paper outlines the challenges in cost estimating knowledge capture. In order to aid the Expert Judgement Rationale capture process a generic inference model is developed for analogy based cost estimating. The model is validated within three industrial sectors: aerospace, automotive and defense. The model identifies the key assumption intensive inferences for the cost estimating. The inferences will provide the foundation for developing a Cost Estimating

Rationale Capture (CERC) tool. Capturing the cost estimating rationale will facilitate concurrent cost estimating through the sharing and reuse of cost estimating knowledge.

References

-
- [1] Hammaker, Joe (2000). *The Faster, Better, Cheaper Approach To Space Missions: A Cost Analysis Perspective*. SSCAG's 69th meeting, European Space Agency, Noordwijk, the Netherlands, May 11-12th.
 - [2] Rush, C., and Roy, R. (2000). *Analysis of cost estimating processes used within a concurrent engineering environment throughout a product life cycle*. 7th ISPE International Conference on Concurrent Engineering: Research and Applications, Lyon, France, July 17th - 20th, Technomic Inc., Pennsylvania USA, 58-67.
 - [3] Rush, C., and Roy, R. (2001). *Capturing quantitative and qualitative knowledge for cost modelling within a concurrent engineering environment*. 8th ISPE International Conference on Concurrent Engineering: Research and Applications, Anaheim, California, July 29th - August 1st, CETEAM International, USA, 209-218.
 - [4] Beltramo, M. N. (1988). 'Beyond Parametrics: The role of subjectivity in cost models.' *Elsevier: Engineering Costs and Production Economics: An International Journal for Industry*, **14**, 131-136.
 - [5] Stensrud, E., and Myrvtveit, I. (1998). *Human performance estimating with analogy and regression models: an empirical validation*. IEEE, Proceedings from the Fifth International Software Metrics Symposium, 205-213.
 - [6] Bashir, H. A., and Thompson, V. (1997). 'Metrics for design projects: a review.' *Design Studies*, **20**(3), 163-277.
 - [7] Hughes, R. T. (1996). 'Expert judgement as an estimating method.' *Information and Software Technology*, **38**, 67-75.
 - [8] Shepperd, M., and Schofield, C. (1997). *Estimating software project effort using analogies*. IEEE Transactions on Software Engineering, **23**(12), 736-743.
 - [9] Lederer, A. L., and Prasad, J. (1998). 'A causal model for software cost estimating error.' *IEEE Transactions on Software Engineering*, **24**(2), 137-148.
 - [10] Dalkey, N., and Helmer, O. (1962). *An Experimental Application of the Delphi Method to the use of Experts*. Contract Number AF 49(683)-700, United States Airforce, RAND Corporation.
 - [11] Dalkey, N., Brown, B., and Cochran, C. (1969). *The Delphi Method, 111: Use of Self Ratings to Improve Group Estimates*. RM-6115-PR, The RAND Corporation.
 - [12] Kitchenham, B. A. (1992). 'Empirical studies of assumptions that underlie software cost-estimation models.' *Information and Software Technology*, **34**(4), 211-218.
 - [13] Rush, C., and Roy, R. (2001). 'Knowledge in Cost Modelling.' *The Cost Engineer: The Journal of the Association of Cost Engineers*, **39**(1), 10-12.
 - [14] Stewart, R. D., Wyskida, R. M., and Johannes, J. D. (1995). *Cost estimators reference manual*, New York: Wiley Interscience.
 - [15] Department Of Defence (1999). *Parametric Estimating Handbook*, 2nd Ed., DoD, <http://www.ispa-cost.org/PEIWeb/cover.htm>.
 - [16] Mileham, R. A., Currie, C. G., Miles, A. W., Bradford, D. T. (1993). A Parametric Approach to Cost Estimating at the Conceptual Stage of Design. *Journal of Engineering Design*, **4**(2), pp. 117-125.

-
- [17] Pugh, P. (1992). 'Working top-down: Cost estimating before development begins.' *Proceedings from the Institution of Mechanical Engineers*, **206**, 143-151.
- [18] Pine, David (2000). *New Ways of Doing Business*. SSCAG's 69th meeting, European Space Agency, Noordwijk, the Netherlands, May 11-12th.
- [19] Boehm, B. W. (1981). *Software engineering economics*, Prentice-Hall: Englewood Cliffs, N.J.
- [20] Shepperd, M., Schofield, C., and Kitchenham, B. (1996). *Effort Estimation Using Analogy*. The 18th International Conference on Software Engineering, Berlin, 170-178.
- [21] Bashir, H. A., and Thompson, V. (2001). 'An analogy-based model for estimating design effort.' *Design Studies*, **22**(2), 157-167.
- [22] Cowderoy, A. J. C., and Jenkins, J. O. (1988). *Cost-estimation by analogy as a good management practice*. Software Engineering 88, Second IEE/BCS Conference, 80 -84.
- [23] Tessem, B., and Modeling, S. (1997). 'Analogy and complex software modelling.' *Computers in Human Behaviour*, **14**(4), 465-486.
- [24] Klein, G., (1998). *Sources of Power: How people make decisions*. Massachusetts Institute of Technology, ISBN: 0-262-11227-2.
- [25] Gentner, D. (1983). 'Structure-mapping: A theoretical framework for analogy.' *Cognitive Science*, **7**(2), 155-170.
- [26] NASA (Accessed July 2001). *Parametric cost estimating process flow analogy approach*. <http://www.jsc.nasa.gov/bu2/analogy.html>
- [27] Zhuge, H., Jian, M., and Xiaoqing, S. (1997). 'Abstraction and analogy in cognitive space: A software process model.' *International Journal of Information and Software Technology*, **39**(7), 463-468.
- [28] Kadoda, G., Cartwright, M., Chen, L., and Shepperd, M. (2000). *Experiences using case-based reasoning to predict software project effort*. Conference on Empirical Assessment in Software Engineering (EASE), Printed by: Keele University, 1-23.
- [29] Duverlie, P., and Castelain, J. M. (1999). 'Cost Estimation During Design Step: Parametric Method Versus Case Based Reasoning.' *The International Journal of Advanced Manufacturing Technology*, **15**, 895-906.
- [30] Rehman, S., and Guenov, M. D. (1998). 'A methodology for modelling manufacturing costs at conceptual design.' *Computers ind. Engng*, **35**(3-4), 623-626.
- [31] Defence Procurement Agency (Accessed July 2001). *Projects: Future Offensive Air Systems (FOAS)*. <http://www.mod.uk/index.php3?page=1496>
- [32] Durkin, J (1994). *Expert Systems, Design and Development*. Macmillian Publishing Company, New York, USA, ISBN: 0-02-330970-9, 1994.
- [33] Jackson, P (1986). *Introduction to Expert Systems*. Addison-Wesley Publishing Company, ISBN: 0-201-14223-6.
- [34] Schreiber, G., Akkermans, H., Anjewierden, De Hoog, R., Shadbolt, N., Van De Velde, W., & Wielinga, B (2000). *Knowledge Engineering and Management: The CommonKADS Methodology*. A Bradford Book, The MIT Press, Cambridge, Massachusetts.
- [35] Bailey, J., Roy, R., Harris, R. And Tanner, A., "Cutting Tool Design Knowledge Capture", *Industrial Knowledge Management - A Micro Level Approach*. Roy, R. (Ed.), Springer-Verlag (London), ISBN 1-85233-339-1, pp. 393-411, 2000.
- [36] Klein, M. (1993). 'Capturing design rationale in concurrent engineering teams.' *Computer*, **26**(1), 39-47.
- [37] Lee, J., and Lai, K.-Y. (1991). 'What's in design rationale?' *Human-Computer Interaction*, **6**(3&4), 251-280.

-
- [38] Conklin, J., and Yakemovic, K. B. (1991). 'A process-oriented approach to design rationale.' *Human-Computer Interaction*, **6**(3&4), 357-391.
- [39] Shum, S. B. (1996). 'Design argumentation as design rationale.' *The Encyclopaedia of Computer Science and Technology*, **35**(20), 95-128.