Information and Software Technology xxx (2011) xxx–xxx

Contents lists available at ScienceDirect

Information and Software Technology

journal homepage: www.elsevier.com/locate/infsof

A method for assessing confidence in requirements analysis

Kenneth Boness [a], Anthony Finkelstein [b], Rachel Harrison [c,*]

[a] School of Systems Engineering, University of Reading, Berks, UK
[b] Dept. of Computer Science, UCL, Gower Street, London WC1E 6BT, UK
[c] Dept. of Computing & Electronics, Oxford Brookes University, Oxford OX33 1HX, UK

ARTICLE INFO

Article history:
Received 24 August 2010
Received in revised form 5 May 2011
Accepted 6 May 2011
Available online xxxx

Keywords:
Requirements analysis
Confidence factors
Meta-level assessment

ABSTRACT

Context: During development managers, analysts and designers often need to know whether enough requirements analysis work has been done and whether or not it is safe to proceed to the design stage.
Objective: This paper describes a new, simple and practical method for assessing our confidence in a set of requirements.
Method: We identified four confidence factors and used a goal oriented framework with a simple ordinal scale to develop a method for assessing confidence. We illustrate the method and show how it has been applied to a real systems development project.
Results: We show how assessing confidence in the requirements could have revealed problems in this project earlier and so saved both time and money.
Conclusion: Our meta-level assessment of requirements provides a practical and pragmatic method that can prove useful to managers, analysts and designers who need to know when sufficient requirements analysis has been performed.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Whenever we attempt to engineer or re-engineer a software system it is widely accepted that arriving at a set of requirements in which we have a lot of confidence is the key to success [20]. However there has been little work to date on ways to arrive at estimates of confidence or on methods that can be used to determine how much confidence different stakeholders attach to a set of requirements. A large system may have very many requirements, each with a different set of confidences associated with it. Estimates of these confidences will help managers to make decisions concerning the costs and benefits of a project.

Previously we have reported on a technique for assessing risks during requirements analysis [2]. During subsequent case studies on real-world systems we came to the realisation that any method intended for use in the real-world has to be as simple and practical as possible if it is to have any hope of being adopted by industry. The new method we describe in this paper uses a simplified form of goal responsibility modelling [27] and replaces the probabilistic risk metrics of [2] with confidence assessments performed by experts using an ordinal scale. This is an important improvement because the probabilistic risk metrics used previously implied a level of precision which could not be guaranteed. Our new method further extends the earlier technique by moderating the assessments using argumentation theory [24] and propagating them within a system using tabulation. Our method is compatible with most requirements representations that depend upon the notion of stepwise refinement. We pay particular attention to the assumptions of stakeholders [8,9,18], which are so often neglected to the detriment of the development.

Our method for assessing confidence during requirements analysis can be summarised as follows:

(1) Construct a goal decomposition graph.
(2) Annotate the graph with estimates of confidence.
(3) Determine the feasibility and adequacy of the requirements.
(4) Consider whether the threats predicted by the feasibility and adequacy assessments are acceptable.

The method is a very practical approach to assessing confidence in requirements which extends our previous technique and is applicable to real-world requirements engineering. Without it managers, analysts, designers and developers are forced to make decisions about whether to continue analysing requirements or start building systems with very little information. The method is particularly useful during requirements analysis and the early stages of systems' development such as the inception and elaboration phases of RUP, and could also be used for planning Scrum sprints. It can be used both with new systems built from scratch and when systems must be constructed using existing COTS

* Corresponding author. Tel.: +44 1865 484506; fax: +44 8700 526595.
E-mail addresses: k.d.boness@reading.ac.uk (K. Boness), a.finkelstein@cs.ucl.ac.uk (A. Finkelstein), rachel.harrison@brookes.ac.uk (R. Harrison).

0950-5849/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.infsof.2011.05.003

components or legacy systems. It can also be used during contract negotiations, or to facilitate responses to invitations to tender. The method has been used during consultancy with a number of our industrial collaborators, and has been found to be helpful.

To demonstrate our method, we present a small but typical problem involving the calculation of body mass index. The method was validated by retrospectively applying it to a case study which we had assessed previously [2]. The case study consisted of a medium-sized project provided by an Administrative Division of University College London. We also carried out semi-structured interviews with the project manager. The results were encouraging and do suggest that our method is useful and usable.

This paper begins with a presentation of our simplified technique for constructing a goal decomposition graph called *goal sketching*. We then go on to discuss the factors which we would like to assess. This is followed by details of our method for assessment and a small but typical exemplar problem. The validation is described next and in the final sections of the paper we discuss related work and our conclusions.

2. Goal sketching

Our assessment method requires the use of a goal graph such as the one shown in Fig. 1. Such a graph could be produced using the KAOS [26] method. However our research, performed with industrial collaborators over many years, convinced us that producing a complete goal graph quickly in a real-world project can be difficult. This led to the development of what we refer to as goal sketching [1,3].

2.1. The goal sketching technique

Goal sketching can be used as a precursor to some other requirements analysis method (such as KAOS modeling [26], use case methods [4], traditional hierarchical requirements modelling [7], etc.) or it can be used alone. It closely resembles KAOS but aims to be very practical. A goal sketch is in fact a goal graph, but the goal sketching technique emphasizes the presence of assumptions and distinguishes them from products (the various system elements to be constructed). This raises awareness of assumptions in the goal analysis [8,9,18,16].

In goal sketching we set down the objectives and show how each objective is to be satisfied (*or at least satisficed [19,23]*) in the resulting system. This is the *keep all objectives satisfied* maxim of van Lamsweerde et al. [26]. There may be many ways to satisfy the objectives, (see, for example, [25]) resulting in the development of a number of different models during analysis. These models allow us to show the project stakeholders' alternative systems for normal operation. Further models can be built for commissioning, decommissioning and maintenance if these are also appropriate.

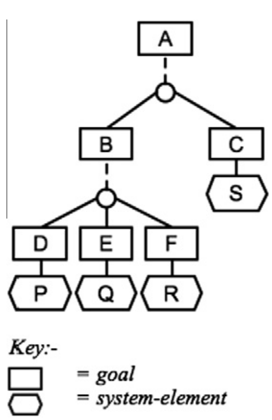

Fig. 1. A simple goal graph.

Fig. 1 illustrates goal-refinement and is referred to as a goal graph. Typically an analyst aiming to specify a system-to-be constructs a hierarchy of goals. The most abstract objectives are presented as root goals from which stem a system of sub-goals refined in steps until goals are reached that are sufficiently concrete that they can be assigned as responsibilities to elements of the intended system. For example, in Fig. 1, *A* is a root goal which is refined into *B* and *C*; both of these are more concrete than *A*. Goal *C* is sufficiently concrete that an element, *S*, of the system-to-be has been given responsibility to guarantee its satisfaction. *B* is not concrete and needs to be refined into *D, E* and *F* which are sufficiently concrete and so have been given to system elements *P*, *Q* and *R* for their satisfaction. This goal graph is said to be structurally complete as every leaf-goal (*C, D, E* and *F*) is guaranteed by an element of the system (*P, Q* and *R*) and consequently, (by refinement arguments joining the leaves to the root), every goal is satisfied.

This state of structural completeness shows how all objectives are understood as being satisfied. When there is uncertainty about an objective or how it might be satisfied it may be necessary for the analyst to approximate some of the goal-refinement in lieu of more complete information. This condition may be found at any point of a project, especially in the early stages such as RUP inception and elaboration. The value of applying the *keep all objectives satisfied* maxim through the device of structural completeness compels the analyst to reveal what is actually known about the requirements so that threats to the stakeholders' expectations can be exposed.

2.2. An example: calculating body mass index

By way of demonstration, we present in Fig. 2 a problem involving the calculation of body mass index that we used in our earlier paper [2].

From the requirements definition in Fig. 2 the goals for the system were determined by the analyst to be as follows:

(1) Normal operation of the walk-on scales must be maintained.
(2) The scales are for use in public places.
(3) WeighCom's good reputation must be maintained.
(4) The scales are to be constructed from prescribed components.

We assume that these goals have been agreed with the stakeholders. We will use the first goal (*normal operation of the walk-on scales must be maintained*) to illustrate our goal sketching technique. The goal sketch for this goal is shown in Fig. 3. This sketch extends the model produced using Objectiver [6] so that responsibility assignments are shown as hexagons. Obstacles are shown as parallelograms, and indicate anti-goals which can prevent a goal from being satisfied. The goals are numbered in the order in which they were analysed. Tool support is needed for both goal sketching and (as discussed later) the application of argumentation. This will ensure that our method is very practical, and is a topic of our current work. This goal sketch will also be used later to illustrate the application of our confidence factors.

3. The confidence factors

In our previous work [2] we identified four independent risk factors: (1) the environmental assumptions, (2) the achievability of the implementation of the requirements, (3) the integrity of the refinements and (4) the stakeholders' mandate. We have

"The customer, WeighCom, wishes to develop new walk-on scales that can be installed in public places and used by any passers-by to measure their weight, height and body mass index (BMI) and receive a business card sized printed record on the spot. Normal operation is for the user to step onto a pressure mat facing an instruction screen and stand under an acoustic ranger. The measurements are made once the user pays a fee of 1 Euro into a receptor.

WeighCom specifies that the solution must use certain components: pressure mat (PM); coin receptor (CR); acoustic ranger (AR) and integrated processor with alpha numerical visual display and user selection touch screen (IP). All of these are to be controlled through software using an API. These components support an existing assembly in which the whole is weather proof and vandal proof.

WeighCom currently installs personal weighing equipment in public places for coin operated use by the public. They have an excellent reputation, which is of paramount importance to them, for always providing a reliable service or repaying. They have a call centre which customers can call if their installations appear to be malfunctioning."

Fig. 2. A requirements definition for the calculation of body mass index.

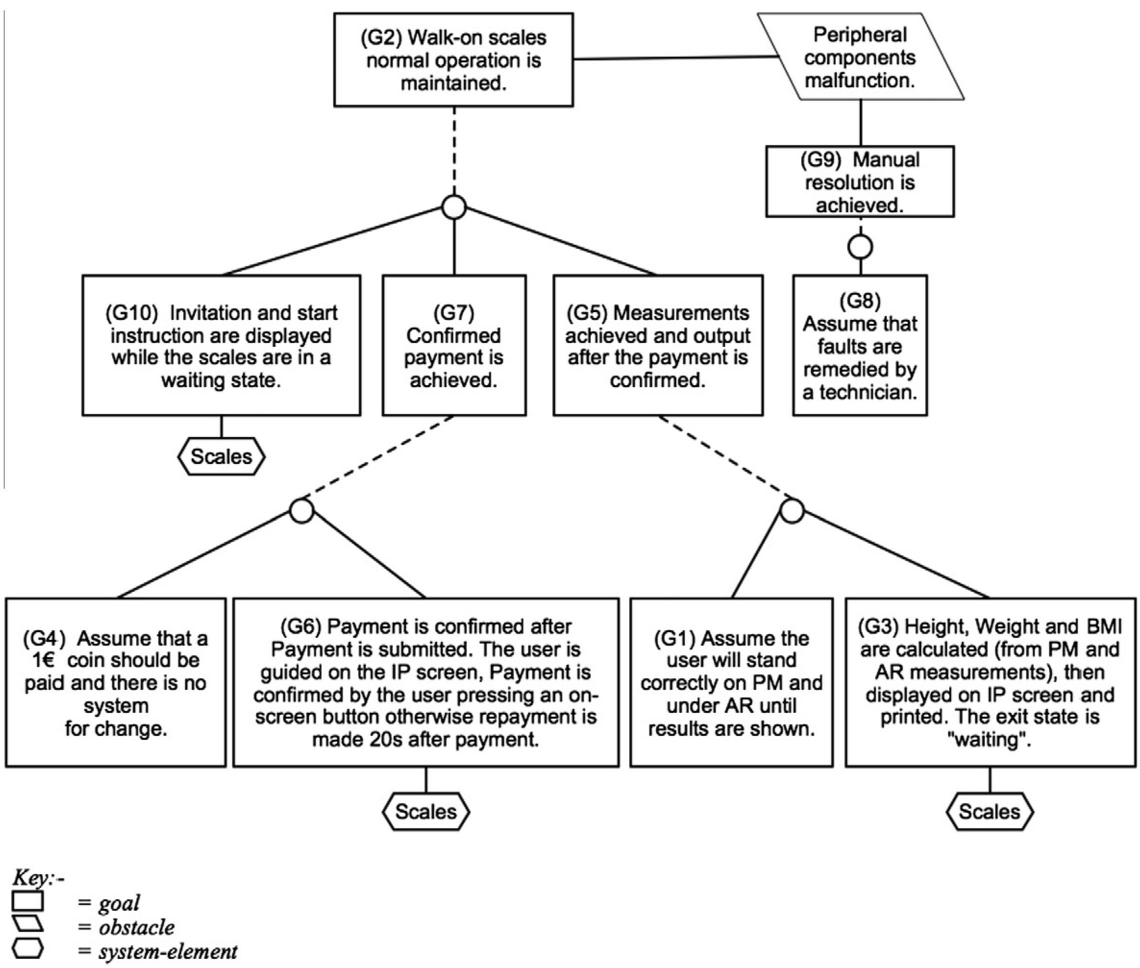

Fig. 3. Goal sketch for the root goal *'Maintain walk-on scales normal operation'* of the requirements definition of Fig. 1 (WeighCom).

clarified and modified these factors slightly in the light of our recent work. The modified factors are described in more detail below.

Our purpose is to appraise what confidence the stakeholders might reasonably place in the analysis expressed in a goal graph. For example, in Fig. 1 the necessary operationalizing elements of the system-to-be (*P, Q, R* and *S*) may not be dependable. Those that are provided by the environment may be infeasible through technical infeasibility or lack of resources. For each operationalizing element which is an assumption (i.e. something we assume we can depend on) we must assess our confidence in its dependability. We call this the *Assumption* confidence factor and denote the raw data for this as ASSUME. For each operationalizing element to be implemented we have a confidence factor called *Achievability*, the raw data for this we denote as ACHIEVE.

Working from the root to the leaves in the goal graph of Fig. 1 we ask whether the root goal (*A*) encompasses the problem to be solved by the system-to-be. This will be determined by our confidence that its sub-goals (*B* and *C*) are sufficient to span the problem and our confidence that the goals *D, E* and *F* adequately satisfy *B*. To gain confidence at each step we rely on the judgements of experts. Thus we require at least one expert's rating of the confidence in the refinement of *A* into *B* and *C* and similarly for the next refinement

Table 1
The four confidence factors.

Confidence factor	Description
Assumption	Confidence that the assumptions are sound
Achievability	Confidence that the acceptance criteria are achievable
Refinement	Confidence that the goal's refinement is sound
Engagement	Confidence that the stakeholders have adequately scrutinised the goal

steps from *B* into *D, E* and *F*. We call this the *Refinement* confidence factor and we denote the raw data provided by the judgement as REFINE. We add to this judgement a separate assessment of the stakeholders' engagement when considering each goal. We call this the *Engagement* confidence factor and denote the raw data provided in this judgement as ENGAGE. The confidence factors are summarised below.

Assumption leaf goals trusted to the environment for satisfaction despite inadequate grounds for believing this is reliable.

Achievability leaf goals assigned to the system's new elements for satisfaction despite inadequate grounds for believing an acceptable implementation is achievable.

Refinement refinements where the stated refinement is open to question (due to semantic entailment), goals with no justifiable parents or where there is an uncertainty about the degree to which a goal contributes to its parent goal, possibly due to errors with the semantic entailment or to 'gold-plating'.

Engagement goals where stakeholder engagement is uncertain, in other words where the stakeholder(s) scrutiny of the goal is in doubt.

Table 1 summarises the descriptions of the confidence factors for ease of reference.

Fig. 1 is a simple goal graph. Goal-graphs produced by KAOS can also have multiple roots, obstacles (anti-goals) and conflicting goals. The goal sketching practice is to collect multiple roots under a single super-ordinate goal so as to ensure all appropriate concerns are satisfied. The *Refinement* factor can serve to ensure that all necessary roots are present. An *obstacle* is a barrier to the proper satisfaction of a goal. In this paper we deal with obstacles as quasi-root goals [2]. There are other alternative treatments (see [3,28]) but these are not discussed here. Similarly in the interests of brevity we do not discuss conflicts save to note that we use assumptions to manage them. In short we have found that the four factors (*Assumption*, *Achievability*, *Refinement* and *Engagement*) are sufficient for our purpose. Clearly the less confidence we have in these factors, the greater is the risk that the project will fail.

Our appraisal method depends on having a goal graph expressing the rationale by which the objectives are expected to be satisfied. To this we append expert judgements at each goal (for *Assumption*, *Achievability*, *Refinement* and *Engagement*) and we use these, as will be explained, as the basis for a Confidence Profile which shows our confidence in the adequacy and the feasibility of the undertaking.

In goal sketching each goal is formed as a proposition about the domain of the system-to-be that must be true for the goal to be satisfied. This formulation allows goal sketching to capture assumptions in the goal graph as each assumption can be expressed as a proposition that must hold in the domain of the system-to-be.

4. Assessing the confidence factors

4.1. The assessment process

Our process for assessing the confidence factors is similar to that of lawyers who must appraise claims which are not true with absolute certainty. For example, lawyers who appraise policy claims have to use the evidence which they are given to support or refute the claim. For each goal we must determine the level of confidence we have with regards to each factor.

Fig. 4 illustrates the theory of argumentation [24] applied to the assessment of confidence factors. Here the *Claim* is the degree of confidence that the assessor has with regards the factor under consideration. The *Data* is the description of the goal and its place in the hierarchy. The *Warrant* is the reasoning behind the claim, based on the data, and the *Backing* records the reasons for taking the warrant seriously. This consists of support such as statistical principles, laws, taxonomies, definitions, standards, best practice, feedback such as records of past performance for the current or related projects, etc. The *Rebuttal* expresses counter-reasoning based on the data [18,20].

The difficulty of providing scales for consistent assessment and interpretation is well known in social, political and other sciences in situations in which it is necessary to use subjective assessments. Guidelines are needed to try to ensure consistent use. For example, a three-point ordinal scale for assessing confidence has been proposed and described in a military context under the heading "What We Mean When We Say: An Explanation of Estimative Language" [21]. We have adapted and extended this to produce the 4-point subjective, comparative scale shown in Table 2. The ratings are mapped onto descriptions of our confidence factors. Thus the confidence factors *Assumption*, *Achievability*, *Refinement* and *Engagement* will be measured on a 4-point ordinal scale (*None*, *Low*, *Medium* and *High*). This scale is applied to all the ratings of claims concerning confidence factors in this paper.

Our process for assigning ratings (from *None* to *High*) follows Fig. 4 for each confidence factor (*Assumption*, *Achievability*, *Refinement* and *Engagement*). The default rating is *Low*. For each confidence factor the Rebuttal challenges the veracity and sufficiency of the data underpinning the Warrant. The Rebuttal thus consists of evidence which attempts to negate the Warrant. The typical constituents of the Claim, Data, Warrant, Backing and Rebuttal for each

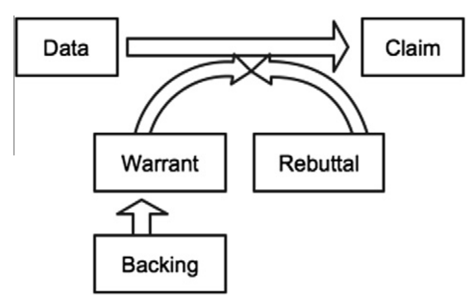

Fig. 4. The structure of arguments in confidence assessment.

Table 2
Comparative ordinal scale for rating claims concerning the confidence factors.

Rating	Description
None	There is a known fact that suggests the claim is unsound
Low	The claim's credibility or plausibility is questionable, or the information is too fragmented or poorly corroborated to make solid analytic inferences, or we have significant concerns or problems with the sources
Medium	The claim is credibly sourced and plausible but not of sufficient quality or corroborated sufficiently to warrant a higher level of confidence
High	The claim is based on high-quality information, and/or the nature of the confidence factor makes it possible to render a solid judgment. A high confidence judgment is not necessarily a certainty, however

Please cite this article in press as: K. Boness et al., A method for assessing confidence in requirements analysis, Inform. Softw. Technol. (2011), doi:10.1016/j.infsof.2011.05.003

of the confidence factors are described and illustrated with examples below.

Confidence factor: Assumption

Claim: The confidence that the assumption can be trusted is [*None*/*Low*/*Medium*/*High*].

Data: Goal description and evidence.

Warrant:

- *None*: there is evidence that the assumption cannot be trusted.
- *Low*: there is no evidence for or against the assumption.
- *Medium*: there is recorded assurance that the assumption is demonstrably sound for the service period of the system.
- *High*: there is recorded and tested assurance that the assumption is demonstrably sound for the service period of the system.

Backing: Evidence of reference to standards [11,12], records of good past performance of the assessor, analyst, and other team members, evidence of attempts to discover underlying assumptions, quality of service guarantees, evidence of adherence to best practice, feedback from ratings of identical or similar trusted assumptions from similar previously implemented systems or prototypes.

Rebuttal: Evidence of lack of reference to standards, records of poor past performance of the assessor, analyst, and other team members, evidence of lack of attempts to discover underlying assumptions, lack of quality of service guarantees, lack of evidence of adherence to best practice, feedback from incorrect ratings of identical or similar trusted assumptions from similar previously implemented systems or prototypes.

Example: Consider Goal G4 of Fig. 3 (*User paid a £1 coin into CR and subsequently pressed a confirm button on the IP screen.*) This is rated as *High* for *Assumption* because during trials of a pilot system 100% of the participants performed this action without prompting.

Confidence factor: Achievability

Claim: The confidence that the goal can be achieved given the project and problem domain constraints is [*None*/*Low*/*Medium*/*High*].

Data: Goal description, project plan and problem domain evidence.

Warrant:

- *None*: there is evidence from the problem domain or project plan that the goal cannot be achieved.
- *Low*: there is no evidence for or against the achievability.
- *Medium*: there is no reason to expect problem domain obstacles and there is at least an outline project plan showing that the goal is achievable within the project constraints.
- *High*: there is no reason to expect problem domain obstacles and there is a project plan showing that the goal is achievable within the project constraints.

Backing: Evidence of reference to standards, recognition of established design patterns, records from good past performance of the assessor, analyst, and other team members, quality of service guarantees, evidence of adherence to best practice, a CMM rating of Level 2 or more, or the equivalent thereof, feedback from *Achievability* ratings of identical or similar previously implemented systems or prototypes.

Rebuttal: Evidence of lack of reference to standards, records from poor past performance of the assessor, analyst, and other team members, lack of quality of service guarantees, lack of evidence of adherence to best practice, a CMM rating of Level 1, or the equivalent thereof, feedback from incorrect ratings of identical or similar previously implemented systems or prototypes.

Example: Consider Goal G10 of Fig. 3 (*An invitation and start instruction to be displayed while the scales are in a waiting state.*) This is rated as *High* for *Achievability* because the company has implemented a similar system in the past year.

Confidence factor: Refinement

Claim: The confidence that the refinement of this goal is sound is [*None*/*Low*/*Medium*/*High*].

Data: Goal description and evidence of sub-goal acceptance criteria, refinement arguments and use cases and activity diagrams for functional goal decompositions.

Warrant:

- *None*: there is evidence that at least one critical sub-goal is missing.
- *Low*: there is no evidence for or against the refinement.
- *Medium*: no acceptance tests are required other than those that apply to the child goals.
- *High*: there is firm agreement that no acceptance tests are required other than those that apply to the child goals. There is also evidence to show that the child goals are individually correct.

Backing: Evidence of reference to standards, recognition of established refinement patterns, feedback from past performance of the assessor, analyst, and other team members, quality of service guarantees, evidence of adherence to best practice, existence of use case and activity diagrams for functional decomposition, feedback from refinement ratings of identical or similar previously implemented systems or prototypes.

Rebuttal: Evidence of lack of reference to standards, feedback from poor past performance of the assessor, analyst, and other team members, lack of quality of service guarantees, lack of evidence of adherence to best practice, lack of existence of use case and activity diagrams for functional decomposition, feedback from incorrect refinement ratings of identical or similar previously implemented systems or prototypes.

Example: Consider Goal G2 of Fig. 3 (*Maintain walk-on scales normal operation.*) This is rated as *High* for *Refinement* because the analyst has successfully worked on a similar system in the recent past.

Confidence factor: Engagement

Claim: The confidence that the stakeholders are engaged with the goal is [*None*/*Low*/*Medium*/*High*].

Data: Goal description and evidence of engagement.

Warrant:

- *None*: there is evidence that the goal is considered inappropriate.
- *Low*: there is no evidence for or against the engagement.
- *Medium*: there is evidence that the goal has been reviewed without objection.
- *High*: there is evidence that the implications of the goal are understood.

Backing: Records from past performance of interaction between the assessor, analyst, and other team members and the stakeholders, evidence of adherence to best practice for stakeholder engagement, evidence of records of meetings between stakeholders and analysts, feedback from engagement ratings of identical or similar previously implemented systems or prototypes.

Rebuttal: Lack of records from past performance of interaction between the assessor, analyst, and other team members and the stakeholders, lack of evidence of adherence to best practice for stakeholder engagement, lack of evidence of records of meetings between stakeholders and analysts, feedback from incorrect engagement ratings of identical or similar previously implemented systems or prototypes.

Example: Consider Goal G1 of Fig. 3 (*User to stand correctly on PM and under AR until results are shown.*) This is rated as *High* for *Engagement* because of evidence that stakeholders have understood this goal. *Engagement* for Goal G7 (*Confirmed payment*) is

Please cite this article in press as: K. Boness et al., A method for assessing confidence in requirements analysis, Inform. Softw. Technol. (2011), doi:10.1016/j.infsof.2011.05.003

rated as *Medium* because records show that the goal was reviewed without objection.

4.2. Assessing feasibility and adequacy

In requirements analysis it is important to determine whether the work is feasible and whether the project will deliver an adequate result [2]. An implementation is *Feasible* if it is achievable and any assumptions associated with the project can be trusted. Thus *Feasibility* considers both whether we can build the system and also whether we can rely on the parts that we have not built. Therefore we estimate *Feasibility*, *F*, by considering the factors *Assumption* and *Achievability*.

An implementation is *Adequate* if it satisfies the goal of the project. To verify this it is necessary to check that (1) the goal has been refined into sub-goals properly and (2) that the stakeholders have engaged with the requirements analysis phase. Thus *Adequacy* asks both whether we are building the system right and also whether we are building the right system. Therefore we estimate *Adequacy*, A, by considering *Refinement* and *Engagement* together. The details of the calculations are explained below.

4.2.1. Assessing feasibility

The *Feasibility* of a goal graph is assessed using the *Assumption* and *Achievability* confidence ratings for the leaves. The assessment of *Feasibility* propagates towards the root from the leaves. Our technique to propagation is that of the *weakest link*. For example, consider the simple goal graph of Fig. 1. Here the *Feasibility* of *A* is the *minimum* of its children's feasibilities (i.e. the minimum of the ratings for *B* and *C*). Thus if confidence in the *Feasibility* of *B* is *Low* and confidence in the *Feasibility* of *C* is *High* then the confidence that *A* is Feasible will be *Low*. Another example is shown in Fig. 5, which shows an abstraction of Fig. 3 (the WeighCom goal sketch). The goal texts have been removed leaving only the numbers (e.g. G2) and the goals G9 and G8 have been omitted in the interests of simplicity. The expert judgements for the *Assumption* and *Achievability* factors are shown in parentheses attached to the leaves of the graph. From these the *Feasibility* metric can be propagated upwards, from the leaves towards the root using the weakest link technique. Thus goal G7 takes a *Feasibility* rating of *Medium* as this is the lowest confidence of goals G4 and G6.

In fact there are two possible approaches to calculating *Feasibility* using the weakest link technique:

1. Use both the *Assumption* and the *Achievability* ratings.
2. Only use the *Achievability* ratings, effectively ignoring the assumptions.

The first approach may be useful in domains in which it is necessary to obtain as much data as possible when making decisions (such as the safety critical domain, for example). Here data concerning the known assumptions may be crucial. The second approach is useful to managers who want to take a decision based solely on the estimated *Achievability*, and who wish to have an expedient and *good enough* estimate of *Feasibility*. This approach will produce a very rough estimate but in some application domains this may be all that is needed. Fig. 5 was produced using approach 1. Applying approach 2 would change the rating for the *Feasibility* of goal G5 to *High*.

If approach 1 is used and assumptions are ignored then we may need to have a separate process for monitoring assumptions such as identifying them and mitigating against them by recording related events and measuring them. These are the *signposting* and *hedging* tactics of Dewar [8,9].

4.2.2. Assessing adequacy

The *Adequacy* of a goal graph is assessed using the *Refinement* and *Engagement* confidence ratings. To calculate *Adequacy* we must combine these and take account of the fact that the *Adequacy* of a goal at some intermediate level in the graph depends upon the accumulation of the estimates of all of its antecedents from the root downwards. Thus an *Adequacy* assessment cannot be based on the combination of one single *Refinement* rating and one single *Engagement* rating because prior to this stage the *Refinement* or *Engagement* may have become questionable. Rather, the *Adequacy* rating must be calculated bearing in mind a possible lack of confidence in the *Refinement* and *Engagement* ratings as the goal tree is traversed from the root towards the leaves.

To accommodate this necessary recursion we proceed as follows. We assume that every goal has been given *Refinement* and *Engagement* ratings by experts on its own merits. Each goal is given a *Raw Adequacy* (*RA*) rating by using its *Refinement* and *Engagement* ratings as indices for Table 3. The result is a confidence rating (from *None* to *High*). This value will be used as the basis for finding the correct *Adequacy* value for the goal.

This is done by repeated lookups using Table 4 as the goal graph is traversed, starting with the root goal and progressing to the leaves. Thus if a goal has a *RA* rating of *Medium*, M, and its parent's *Adequacy* is *Low*, L, then its resulting *Adequacy* rating, called the *Stepwise Adequacy*, will also be *Low*, and so on.There are two rules which deal with special cases when deciding a goal's *Stepwise Adequacy* from Table 4:

1. If the goal is a root use the row with a *High RA* rating.
2. If the goal has multiple parents use the parents' worst *Adequacy* value.

4.2.3. The Confidence Profile

Once the *Adequacy* and *Feasibility* ratings have been propagated through the goal graph we can then use the total cost of the project

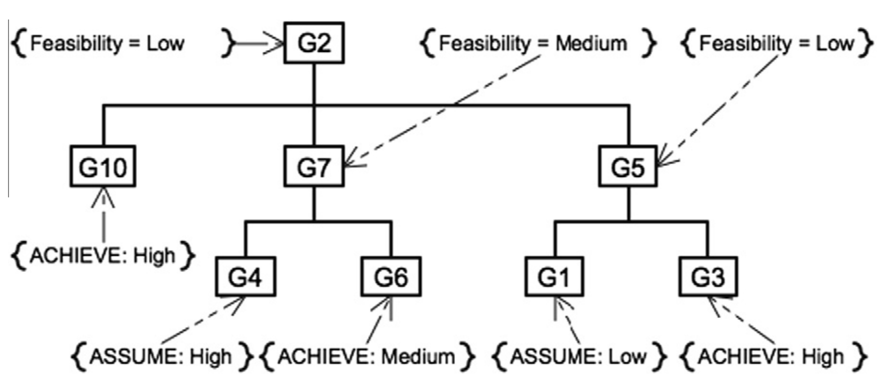

Fig. 5. The calculation of Feasibility for a sub-graph of Fig. 3.

Please cite this article in press as: K. Boness et al., A method for assessing confidence in requirements analysis, Inform. Softw. Technol. (2011), doi:10.1016/j.infsof.2011.05.003

Table 3
Raw Adequacy.

Engage				
H	N	L	M	H
M	N	L	M	M
L	N	L	M	M
N	N	N	N	N
	N	**L**	**M**	**H**
	Refine			

Key: N = *None*; L = *Low*; M = *Medium*; H = *High*.

Table 4
Stepwise adequacy lookup.

Adequacy of Parent				
H	N	L	M	H
M	N	L	M	M
L	N	L	L	L
N	N	N	N	N
	N	**L**	**M**	**H**
	RA			

Key: N = *None*; L = *Low*; M = *Medium*; H = *High*.

to calculate the fractional cost of goals with each rating. There are many ways that we can represent these assessments, which are part of the Confidence Profile for the system. One way is to sum the costs of goals with each level of confidence and tabulate the results. Table 5 shows a typical (contrived) example of this. The data in the cells of the table are aggregated values from the goals. The table highlights the proportion of the project's costs that arise from goals that are assessed as inadequate or infeasible. Here 70% of the project's costs arise from goals that have been assessed with both *None* or *Low Adequacy* and also *None* or *Low Feasibility*, and so are in the *Do Not Proceed* (or *Danger*) zone. The higher the proportion of goals that are in the *Do Not Proceed* zone the greater is the risk to the project. For example, if confidence in a safety critical project is assessed with 70% (say) of the goals in the *Do Not Proceed* zone then this would be a cause for great concern. Measures other than cost are also possible such as sales importance, architectural importance and value to the customer.

5. Example: calculating body mass index

5.1. The annotated goal sketch

The goal sketch for the WeighCom system (Fig. 3) was annotated with confidence factors given by expert judgement (see Fig. 6). Table 6 shows how to present the Confidence Profile on a per-goal basis using the raw values of the confidence factors taken from the annotated goal sketch of Fig. 6 together with estimates of cost and value.

5.2. Assessing feasibility for the WeighCom system

The simplified goal sketch for the WeighCom system shown in Fig. 5 shows the expert ratings of ASSUME and ACHIEVE at the leaves and the weakest link (i.e. the worst case) propagation of *Feasibility* working up from the leaves to the root, as discussed in Section 4.2.1.

5.3. Assessing adequacy for the WeighCom system

Adequacy is assessed using Tables 3 and 4 together with the raw values of the confidence factors taken from the goal sketch of Fig. 6, leading to the results in Table 7.

The metrics *Feasibility* (*F*) and *Adequacy* (*A*) were calculated using the techniques described in Section 4.2 and are shown in Table 8 together with proportional cost (%Cost) and proportional value (%Value).

The Confidence Profile is shown in tabulated style in Fig. 7. The data in the cells of the tables are the aggregated values from the goals. The Goal Count of the rightmost chart in Fig. 7 shows that four leaf goals have *Proceed* status. From the other charts we can see that this amounts to 75% of the proportional value and 70% of the proportional cost. One goal amounting to 30% of the cost and 13% of the value has *Caution* status and one goal with no cost (an assumption) but representing 12% of the value is at a *Do not proceed* status.

6. Validation

To validate our new method we retrospectively applied it to the same case study that we used previously, the University Personal Identifier (UPI) upgrade project [2]. This is a project that took place at University College London (UCL). It was given the go-ahead in 2003 on the basis of a PRINCE2-style Project Initiation Document (PID) and was scheduled to finish in June 2004. In the event it was twice re-planned before it concluded to the satisfaction to the stakeholders in 2006. This project provided a rich set of project data, including the necessary requirements data as well as accompanying historical documentary evidence needed to check the results. In our earlier work we used quantitative probabilistic assessment of risks, and so produced a numerical risk assessment with the help of the project manager and project supervisor. Consequently for the validation of our new method we had to map the original numerical assessments to our new 4-point scale shown in Table 2 before applying the new method. We then compared the results with those obtained previously. We also carried out semi-structured interviews with the project manager to find out

Table 5
Contrived Confidence Profile showing typical fractional costs of goals with particular *Adequacy* and *Feasibility* ratings.

Adequacy				
H	10%	5%	5%	5%
M	5%	5%		5%
L	5%	10%	5%	
N	10%	10%	10%	10%
	N	L	M	H
	Feasibility			

Key

(green)	Proceed (Adequacy = M ∨ H) ∧ (Feasibility = M ∨ H)
(amber)	Caution ((Adequacy = L) ∧ (Feasibility = M ∨ H)) ∨ ((Feasibility = L) ∧ (Adequacy = M ∨ H))
(red)	Do Not Proceed (Adequacy = N ∨ L) ∧ (Feasibility = N ∨ L)

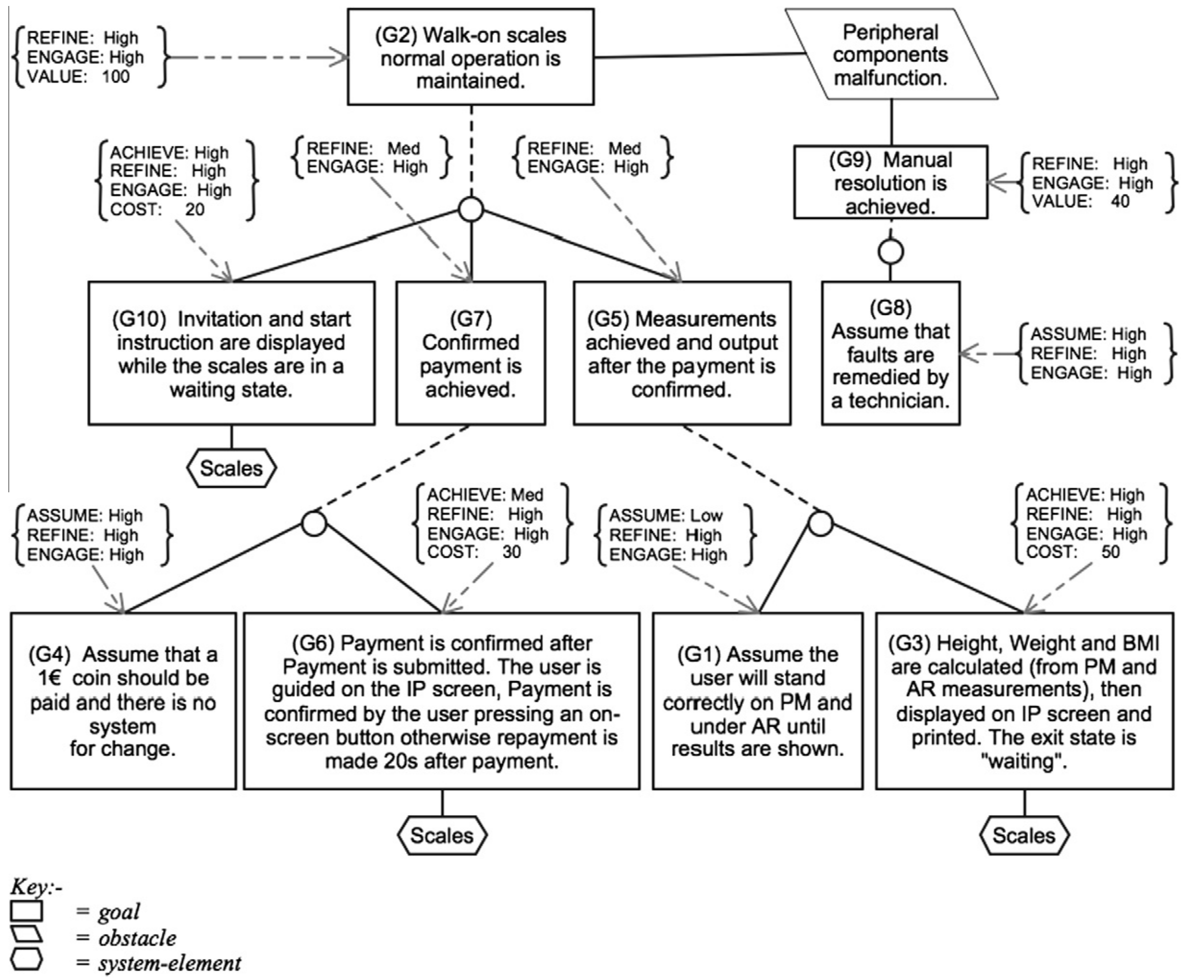

Fig. 6. WeighCom normal operation annotated goal sketch showing raw data.

Table 6
Confidence Profile for WeighCom, root goal 'Maintain walk-on scales normal operation'.

Name	ASSUME	ACHIEVE	REFINE	ENGAGE	Cost	Value
G1	Low		High	High	0	17
G2			High	High	100	100
G3		High	High	High	50	18
G4	High		High	High	0	17
G5			Medium	High	50	35
G6		Medium	High	High	30	18
G7			Medium	High	30	35
G8	High		High	High	0	40
G9			High	High	0	40
G10		High	High	High	20	30

Table 7
Adequacy for the WeighCom system.

Name	REFINE	ENGAGE	Raw adequacy	Stepwise adequacy
G1	High	High	High	Medium
G2	High	High	High	High
G3	High	High	High	Medium
G4	High	High	High	Medium
G5	Medium	High	Medium	Medium
G6	High	High	High	Medium
G7	Medium	High	Medium	Medium
G8	High	High	High	High
G9	High	High	High	High
G10	High	High	High	High

Table 8
Confidence Profile for WeighCom leaf goals.

Goal	Feasibility	Adequacy	%Cost	%Value
G1	Low	Medium	0	12
G3	High	Medium	50	13
G4	High	Medium	0	12
G6	Medium	Medium	30	13
G8	High	High	0	29
G10	High	High	20	21
			100	100

whether or not our new assessments were accurate, and how successful the project was felt to be one year after completion. The project manager was independent of our research and had no vested interest in the results. The following sub-sections introduce the project briefly and then describe the validation.

6.1. Background

The College had been using the UPI system to control access to services over the College intranet. This UPI system links information held in a number of sub-systems. These included registration and human resource sub-systems as well as library sub-systems, buildings access sub-systems etc. Some of the sub-systems were stable, some were unstable and others were under development.

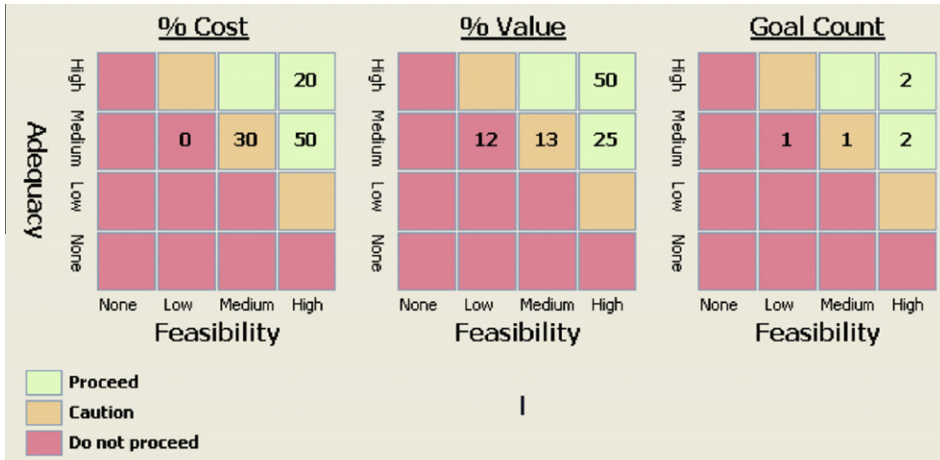

Fig. 7. Confidence Profile for WeighCom.

The existing UPI system was deemed unsatisfactory and needed to be replaced. The idea was to upgrade the existing system to produce a maintainable and efficient replacement that would ensure that staff and students (and others) have appropriate access to the services and facilities of the college (library, buildings, management data and networked software, etc.) and would facilitate person-centric analyses of information available on the intranet. The upgrade project was performed as an in-house reengineering of the system. Success depended upon co-operation with various departments of the college and consequently the project was a hostage to frequent changes emanating from these departments. As the study was conducted on historical data we could compare outcomes with predictions. Using historical data and semi-structured interviews allowed us to determine whether the Confidence Profile and predictions were accurate.

A partial goal sketch for the UPI project is shown in Fig. 8. This shows only the root segment of the entire goal sketch; the remainder of the graph is represented by the dashed lines attached to the lower goals. The whole graph has 52 goals, 32 of which are leaf goals. The graph has six levels at its deepest parts. The tags ($G\langle n\rangle$) are unique goal identifiers produced by our prototype goal sketching tool.

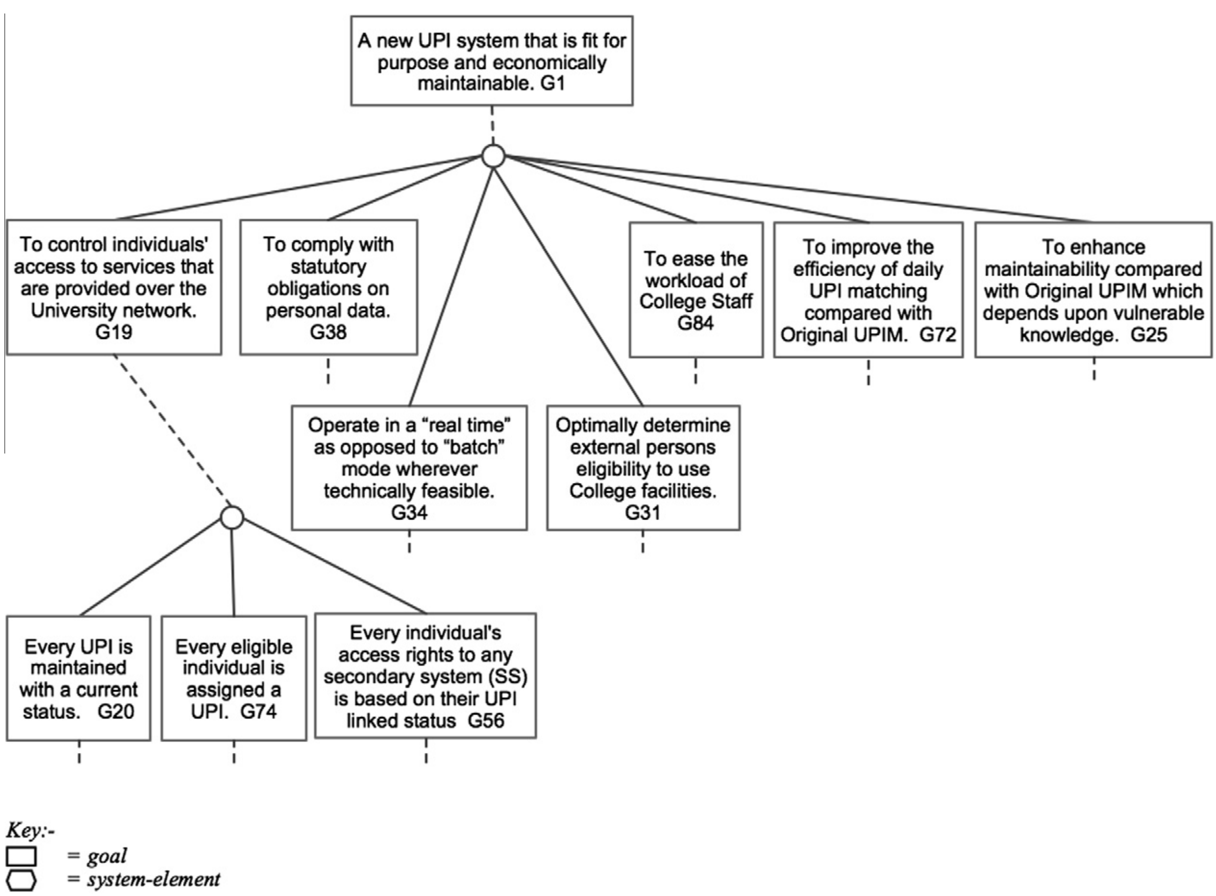

Fig. 8. Root goal sketch for the UPI Project.

Please cite this article in press as: K. Boness et al., A method for assessing confidence in requirements analysis, Inform. Softw. Technol. (2011), doi:10.1016/j.infsof.2011.05.003

ARTICLE IN PRESS

Table 9
Mapping risk factors to confidence factors.

Original risk factor ratings for x	Confidence factor ratings for *x*
$0.7 \leqslant x \leqslant 1.0$	High
$0.4 \leqslant x < 0.7$	Medium
$0.1 \leqslant x < 0.4$	Low
$0 \leqslant x < 0.1$	None

6.2. Applying our new method to the UPI project

The *Assumption* and *Achievability* ratings were derived from the original risk factor ratings given by the project manager [2] by applying the mapping shown in Table 9.

The *Refinement* and *Engagement* ratings were also obtained by transforming the original risk factor ratings. The new tabulation technique for propagating *Adequacy* was then applied. The Confidence Profile is shown in the tables of Fig. 9 and predictions in Table 10. The goal numbering in Table 10 has been kept the same as in the previous study ([2]) for ease of reference.

The *Do Not Proceed* zone in Fig. 9 includes all cells with a rating of *None* (for either *Feasibility* or *Adequacy*) as well as all cells that are rated both *Medium* and *Low*, since a *High* confidence rating for *Adequacy* (say) does not mean that an accompanying rating of *None* for *Feasibility* is acceptable.

The purpose of the metrics is to predict potential problems, or *pain*. In [2] we estimated the threat to the project by calculating the *threat ratios*, where a threat ratio is the ratio of the number of leaf goals (contributing to each of the root goals) that are in the *Do Not Proceed* zone to the total number of leaf goals. This is just one of several possible predictive devices. The pain predictions are calculated from the threat ratio: a ratio greater than or equal to 0.66 is classed as a *High* pain prediction, a ratio greater than or equal to 0.33 but less than 0.66 is classed as a *Medium* pain prediction and a ratio greater than or equal to zero but less than 0.33 is classed as a *Low* pain prediction. The bold face predictions in Table 10 indicate predictions that do not match the experienced pain, whereas italics indicate untested predictions. Goal 31 is missing from the *Experienced Pain* column of Table 10 because it was removed very early on in the project as it became evident that it was inappropriate.

6.3. Mapping the original numerical assessment to our new scales

To perform the mapping, the original *Adequacy* and *Feasibility* values for the leaves of the graph were converted from their

Table 10
New Method Prediction Results.

Goal	Experienced Pain	Threat ratio	Pain prediction
20	Low	2/7	Low
25	Low	0/2	Low
31		1/1	*High*
34	Med	3/5	Med
38	Low	0/2	Low
56	Med	5/7	**High**
72	High	2/2	High
74	Low	0/8	Low
84	High	2/4	**Med**

Key: Bold: Incorrect prediction.
Italic: Untested prediction.

probability ratings to an ordinal scale as shown in Table 9. Applying this scale to the results reported in [2] gives us Fig. 10 and Table 11.

Table 11 shows the actual experienced pain for each of the root goals, the threat ratio and pain predicted in the original work, together with the threat ratio and pain predictions produced after mapping the quantitative values to the new subjective scale of Table 3. The middle column (*Pain predicted in* [2]) shows that in [2] the four predictions in plain text were correct whereas the four in bold did not match the experienced pain.

In Table 11 the results in the rightmost column (*Mapped pain prediction*) show an improvement with six correct and two incorrect predictions. This improvement is due to the enlargement of the *Do Not Proceed* region used in [2].

6.4. Comparing the results obtained with those obtained previously

From Tables 10 and 11 we can see that the pain predictions (in the final columns of each table) are the same except for Goal 72, which is now correctly predicted to result in *High* pain (see Table 10) and Goal 84, which is now incorrectly predicted to result in *Medium* pain. Using the new method the pain predictions are correct for 6 goals and incorrect for 2 goals, whereas previously 4 of the 8 predictions were incorrect in [2], as shown by the middle column of Table 11. The results are therefore more accurate using the new method, and the assessment effort was greatly reduced. The threat ratio and predictions of pain were greatly simplified by using the new method.

The goal sketching method needed only two one-hour face-to-face sessions with the project manager to establish the goal graphs and expert judgements. This represents a very small proportion of the time that had been spent by the project manager on the Project

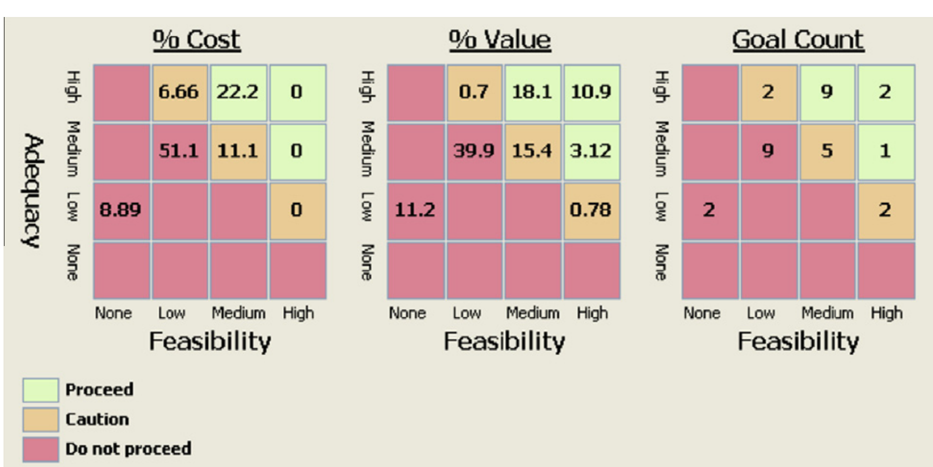

Fig. 9. Confidence Profile for the UPI project using the new method.

Please cite this article in press as: K. Boness et al., A method for assessing confidence in requirements analysis, Inform. Softw. Technol. (2011), doi:10.1016/j.infsof.2011.05.003

ARTICLE IN PRESS

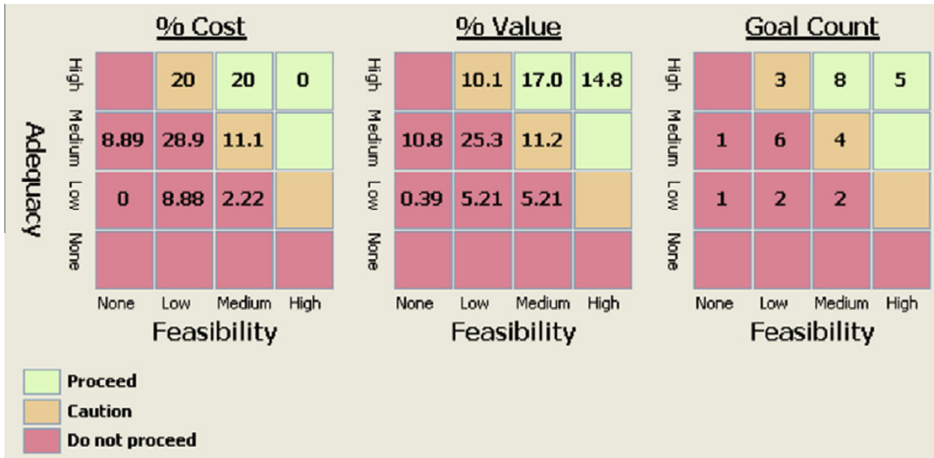

Fig. 10. Original results for the UPI Project.

Table 11
Original Results.

Goal	Experienced pain	Threat ratio in [2]	Pain predicted in [2]	Mapped threat ratio	Mapped pain prediction
20	Low	2/7	Low	2/7	Low
25	Low	0/2	Low	0/2	Low
31		1/1	*High*	1/1	*High*
34	Med	5/5	**High**	3/5	Med
38	Low	1/2	**Med**	0/2	Low
56	Med	7/7	**High**	5/7	**High**
72	High	0/2	**Low**	1/2	**Med**
74	Low	0/8	Low	0/8	Low
84	High	4/4	High	4/4	High

Key: Bold: Incorrect prediction.
Italic: Untested prediction.

Initiation Document alone. Thus the amount of effort that would have been needed to establish and maintain the goal graph and profiles during the project would have been small compared with other project management tasks.

By helping with the construction of the goal graph the project manager was easily able to retrospectively reconstruct the project rationale. It is reasonable to expect that doing this at the start of the project would have provided useful insights about the necessary and sufficient set of sub-systems and critical assumptions.

7. Related work

A great deal of work has been reported in the literature on requirements engineering in general and requirements analysis in particular. The work discussed in this section is specifically related to the assessment of confidence during requirements analysis, i.e. the assessment of knowledge concerning a set of requirements during early stages of the project lifecycle.

The Defect Detection and Prevention (DDP) process was developed at JPL to facilitate risk management during the project life cycle [5] for aerospace systems. It uses a set of risk elements and trees of requirements. The process is intended for use at the architecture stage, and uses a large number of metrics together with pre-determined industry-specific knowledge. This contrasts with our method, which we have deliberately kept as simple and lightweight as possible.

Ruhe et al. have presented an approach for use in requirements negotiation called "Quantitative WinWin" [22], which adds numerical assessments to Boehm's original WinWin work. The approach uses the Analytical Hierarchy Process for a stepwise determination of stakeholders' preferences which is then combined with early effort estimates to evaluate the feasibility of alternative requirements in terms of implementation effort. This work is concerned with finding subsets of requirements that can be implemented without exceeding the maximum effort available whereas our work is concerned with assessing the potential for building the wrong system.

Marchant et al. describe an investigation into how a metric can be applied to requirements gathering to determine the likely success of re-engineering legacy systems [17]. The proposed confidence metric was used in two industrial case studies to analyse the probable success of the projects. The metric is for use during the re-engineering of legacy systems, whereas our method is intended for use during the requirements analysis stage of any project.

Knauss et al. describe an empirical study into assessing software requirements specifications quality [15]. They analysed 40 projects developed by undergraduate software engineering students using a quality model for software requirements. Their results suggest that there is a relationship between the quality of the requirements engineering undertaken and project success. However, this work is directed at assessing software requirements specifications themselves, whereas our work is at a meta-level: we are assessing what is known about the software requirements.

Problem Oriented Software Engineering (POSE) [10] is also concerned with the adequacy of requirements. The meta-analysis of requirements used in our method could also be used with POSE. Similarly the goal-structuring notation (GSN) proposed by Kelly [13,14] involves goals supplemented by contexts and justifications that are refined in a stepwise manner. The results from a GSN analysis could be abstracted into a goal sketch to which our confidence factors could be applied. These ideas will be investigated in future work.

Please cite this article in press as: K. Boness et al., A method for assessing confidence in requirements analysis, Inform. Softw. Technol. (2011), doi:10.1016/j.infsof.2011.05.003

8. Threats to validity

The validity of this study is limited by a number of considerations [29]. Here we consider the threats to construct, internal and external validity.

8.1. Construct validity

Construct validity is concerned with whether the variables used in the study accurately measure the concepts they purport to measure, i.e. has the data that we collected given us accurate estimates of *Assumptions*, *Achievability*, *Refinement* and *Engagement*. We identified two possible threats to construct validity. Firstly, the mapping from risk factor probabilities to discrete confidence factor ratings is somewhat subjective and could have been performed differently, although we believe that our technique is reasonable and defensible. Secondly, during the interviews held with the project manager it became clear that subtly different interpretations of the *Refinement* confidence factor are possible. However, we clarified our description of this factor and also one of the authors worked closely with the project manager to try to ensure that this threat was minimized through shared understanding.

8.2. Internal validity

Internal validity in this study is concerned with whether the relationship between the confidence factors and the threats to the project is a causal relationship, and has not simply arisen as a result of chance. We are reliant on the assessments reported in [2], which were performed over a period of time (several months) and this could be a threat to validity because the assessors may have changed their criteria in this time. However, the assessments were discussed at length by the authors and also with the project manager, both in the first study and during the study reported here.

8.3. External validity

External validity is concerned with the generalizability of the results. The fact that we retrospectively used one medium-sized case-study during the validation of the new method is clearly a threat to external validity. However, this project is typical of the re-engineering of data-intensive systems which can run into problems because of a lack of attention to feasibility and adequacy.

9. Future work

Further work remains to be done on validation. This includes the application of our method to a range of systems from different application domains via collaboration with our industrial partners to determine the application domains to which the method is best suited. This will help to strengthen the validation of the method and may lead to improvements for the real-world application of the method. In particular we wish to improve our guidelines for reviewers as to the use of our 4-point rating scale for assessing the confidence factors and investigate the use of rebuttals so that this part of our method can become accepted as best practice. We have an on-going agenda which is to address the issue of how analysts, designers and developers can move between the stages of software development with confidence. Identifying the risks posed by incomplete requirements and reasoning about them will help project managers working under time pressures make informed decisions about whether to continue or stop requirements engineering work prior to taking design decisions. In the fullness of time we hope to provide an automatable method to facilitate requirements engineering. There is still much to be done before this can be achieved.

The method presented in this paper could also be applied to system testing. Testing is an important part of validating systems. A necessary condition for test completeness is that there is confidence in the refinement argument (which is measured by our *Refinement* factor). A second necessary condition is that every leaf goal has a test case (this can be measured by our *Assumption* and *Achievability* factors). In this way a manager with limited resources can determine how to apportion the resources for maximum benefit. This is something we will explore further in future work.

10. Conclusions

This paper has described a new method for assessing confidence during requirements analysis. We identified 4 factors which can cause difficulties (*Assumptions*, *Achievability*, *Refinement* and *Engagement*). We described a method to assess these factors and from this to estimate the *Feasibility* and *Adequacy* of the proposed work. The method extends our earlier work by using a 4-point ordinal scale, moderating the expert assessments using argumentation theory and propagating the assessments using tabulation. The method was illustrated with a small but typical problem and was validated by retrospectively applying it to a case study which we had assessed earlier and comparing the results. Our new method is thus a refinement and extension of the original technique, and was found to be much more straightforward to apply in practice. The method is being used during consultancy with a number of our industrial collaborators, and has been found to be useful. Our meta-level assessment of a system's requirements provides a practical and pragmatic method that will prove useful to managers, analysts and designers who need to know whether enough requirements analysis work has been done and whether or not it is safe to proceed to the design stage. Future work will include further validation to determine the application domains to which this method is best suited, and the construction of tools to fully automate the method.

Acknowledgments

The authors would like to thank their collaborators for providing project data, particularly Kathryn Lewis of MSD, and the anonymous referees who reviewed this paper and provided very helpful comments.

References

[1] K. Boness, R. Harrison, Goal sketching: towards agile requirements engineering, in: International Conference on Software Engineering Advances, 2007, ICSEA, 2007, pp. 71–76.
[2] K. Boness, A. Finkelstein, R. Harrison, A lightweight technique for assessing risks in requirements analysis, IET Software 2 (1) (2008) 46–57.
[3] K. Boness, Goal-Refinement Appraisal of Soft-Projects (GRASP), Ph.D. Dissertation, University of Reading, 2011.
[4] A. Cockburn, Writing Effective Use Cases, Addison-Wesley, Reading, MA, 2002.
[5] S.L. Cornford, M.S Feather, K.A. Hicks, DDP – a tool for life-cycle risk management, in: Proceedings of the IEEE Aerospace Conference, Big Sky, Montana, 2001, pp. 441–451.
[6] R. Darimont, Requirements engineering with objectiver: from goal analysis to automatically derived requirements documents, in: RE'03 Exhibitors' Track, International Conference on Requirements Engineering (RE03), IEEE Computer Society Press, Los Alamitos, California, 2003.
[7] A.M. Davis, Just Enough Requirements Management, Dorset House Publishing, 2005.
[8] J.A. Dewar, M.H. Levin, Assumption-Based Planning for Army 21. RAND CORPORATION-REPORTS-R-ALL SERIES-.4172, 1992.
[9] J.A. Dewar, Assumption-based Planning: a Tool for Reducing Avoidable Surprises, Cambridge Univ. Press, 2002.
[10] J.G. Hall, L. Rapanotti, M. Jackson, Problem oriented software engineering: a design-theoretic framework for software engineering, in: Fifth IEEE

International Conference on Software Engineering and Formal Methods, 2007, SEFM, 2007, pp. 15–24.

[11] IEEE, IEEE Guide for Developing System Requirements Specifications, IEEE Std 1233, 1998 ed., 1998 doi:10.1109/IEEESTD.1998.88826.

[12] IEEE, IEEE Recommended Practice for Software Requirements Specifications, IEEE Std 830-1998, 1998, doi:10.1109/IEEESTD.1998.88286.

[13] T.P. Kelly, J. McDermid, Safety case construction and reuse using patterns, in: 16th International Conference on Computer Safety and Reliability (SAFECOMP'97), York, 1997.

[14] T.P. Kelly, The Goal Structuring Notation (GSN). Ph.D. Thesis, 1998.

[15] E. Knauss, C. Boustani, T. Flohr, Investigating the impact of software requirements specification quality on project success, in: Proceedings 10th International Conference on Product-Focused Software Process Improvement, PROFES 2009, Oulu, Finland, June 15—17, 2009. Lecture Notes in Business Information Processing, vol. 32, 2009, doi:10.1007/978-3-642-02152-7, 2009, Springer.

[16] M.M. Lehman, The role and impact of assumptions in software development, maintenance and evolution, in: IEEE International Workshop on Software Evolvability, 2005, pp. 3–16.

[17] J. Marchant, C. Tjortjis, M. Turega, A metric of confidence in requirements gathered from legacy systems: two industrial case studies, in: Conference on Software Maintenance and Reengineering (CSMR'06), 2006, pp. 355–361.

[18] R.O. Mason, I.I. Mitroff, Challenging Strategic Planning Assumptions: Theory, Cases, and Techniques, John Wiley & Sons Inc., 1981.

[19] J. Mylopoulos, L. Chung, B. Nixon, Representing and using nonfunctional requirements: a process oriented approach, IEEE Transactions on Software Engineering 18 (6) (1992) 483–497.

[20] B. Nuseibeh, S. Easterbrook, Requirements engineering: a roadmap, in: Proceedings of the Conference on the Future of Software Engineering, Limerick, Ireland, June 04–11, 2000. ICSE '00. ACM, New York, NY, 2000, pp. 35–46. <http://doi.acm.org/10.1145/336512.336523>.

[21] ODNI, Iran: Nuclear Intentions and Capabilities. oai.dtic.mil, 2007.

[22] G. Ruhe, A. Eberlein, D. Pfahl, Quantitative WinWin: a new method for decision support in requirements negotiation, in: Proceedings of the 14th International Conference on Software Engineering and Knowledge Engineering (SEKE '02), Ischia, Italy, July 2002, pp. 159–166.

[23] H.A. Simon, The Science of Management Decisions, 1960.

[24] S.E. Toulmin, The Uses of Argument, Cambridge University Press, 1958.

[25] J.R. Turner, R.A. Cochrane, Goals-and-methods matrix: coping with projects with ill defined goals and/or methods of achieving them, International Journal of Project Management 11 (1993) 93–102.

[26] van Lamsweerde, A. Dardenne, B. Delcourt, F. Dubisy, The KAOS project: knowledge acquisition in automated specification of software, in: Proceedings of the AAAI Spring Symposium Series, Stanford University, AAAI, March 1991, pp. 59–62.

[27] van Lamsweerde, Goal-oriented requirements engineering: a roundtrip from research to practice, in: 12th IEEE International Requirements Engineering Conference (RE'04), 2004, pp. 4–7.

[28] A. van Lamsweerde, E. Letier, Handling obstacles in goal-oriented requirements engineering, IEEE Transactions on Software Engineering, Special Issue on Exception Handling 26 (10) (2000) 978–1005.

[29] C. Wohlin, P. Runeson, M. Höst, M.C. Ohlsson, B. Regnell, Experimentation in Software Engineering: An Introduction, Kluwer Academic Publishers, Norwell, MA, USA, 2000.

Please cite this article in press as: K. Boness et al., A method for assessing confidence in requirements analysis, Inform. Softw. Technol. (2011), doi:10.1016/j.infsof.2011.05.003