Facilitating Use of Assurance Cases in Industries by Workshops with an Agent-Based Method

Yutaka MATSUNO†, Toshinori TAKAI††, and Shuichiro YAMAMOTO†††, Members

SUMMARY Assurance cases are documents for arguing that systems satisfy required properties such as safety and security in the given environment based on sufficient evidence. As systems become complex and networked, the importance of assurance cases has become significant. However, we observe that creating assurance cases has some essential difficulties, and unfortunately it seems that assurance cases have not been widely used in industries. For this problem, we have been developing assurance cases creation methods and opening workshops based on the creation methods. This paper presents an assurance cases creation method called “D-Case Steps” which is based on d* framework [1], an agent-based assurance case method, and reports the results of workshops. The results indicate that our workshops have been improved and our activities on assurance cases facilitates use of them in Japan. This paper is an extended version of [2]. We add detailed background and related works, workshops results and evaluation, and lessons learned from our a decade experiences.

key words: assurance cases, GSN (goal structuring notation), workshop

1. Introduction

Safety cases (assurance cases for system safety) are required to be submitted to certification bodies for developing and operating safety critical systems, e.g., automotive, railway, and defense systems, as well as nuclear power plants and off-shore drilling platforms. There are several standards, e.g. EUROCONTROL [3], the Rail Yellow Book [4], and the MoD Defense Standard 00-56, which mandate the use of safety cases. There are several definitions for assurance cases. We show one such definitions as follows: “A structured argument, supported by a body of evidence that provides a compelling, comprehensible and valid case that a system is safe for a given application in a given environment” [5].

Figure 1 shows the basic structure of assurance cases. To clarify the structure of assurance cases, several graphical notations have been proposed such as GSN (Goal Structuring Notation) [6] and CAE (Claim, Argument, Evidence) [5]. This paper uses GSN.

We briefly explain constructs and their meanings in GSN. Arguments in GSN are structured as trees with a few kinds of nodes, including: goal nodes for claims to be argued for, strategy nodes for reasoning steps that decompose a goal into sub goals, and evidence nodes for references to direct evidence that respective goals hold. Figure 2 is a simple example of GSN. The GSN diagram is a simplified version of an example in the GSN Community Standard [7], which argues safety of a control system.

The root of the tree must be a goal node, called the top goal, which is the claim to be argued for (G_1 in Fig. 2). For G_1, a context node C_1 is attached to complement. Context nodes are used to describe the context (environment) of the goal attached to. A goal node is decomposed through a strategy node (S_1) into sub goal nodes (G_2 and G_3). The strategy node contains an explanation, or reason, for why the goal is achieved when the sub goals are achieved. S_1 explains the method used when arguing (arguments over each
possible fault: A and B. When successive decompositions reach a sub goal \((G, 2)\) that has a direct evidence of success, an evidence node \((E, 1)\) referring to the evidence is added. Here we use a result of fault tree analysis (FTA) as the evidence. For the sub goal \((G, 3)\) that is not decomposed nor supported by evidence, a node (a diamond) of type undeveloped is attached to highlight the incomplete status of the case.

However, even using such graphical notations, creating assurance cases is difficult. System attributes such as safety and dependability of a system involve almost all aspects of the system, and the argument structure tends to be diverge and huge. Also, as the nature of goal oriented structure, decomposing a goal into sub-goals has essential difficulties. Some of the difficulties are indicated in [8] as follows.

1. Mixing up of strategies and goals.
2. Content that should be written as a claim being expressed in the form of an action or function statement rather than as a proposition.
3. Misunderstanding of strategies as judgment branches.
4. Decomposing Into Function Execution Sequences Instead of Arguments.

The work [8] conducted experiments of creating assurance cases by Japanese engineers, and argued that assurance case patterns such as architectural decomposition is effective for preventing such difficulties.

Based on these observations, we have been developing assurance cases creation methods and opening workshops for a decade [9]. Currently we use an assurance case creation method called “D-Case Steps”, which is based on \(d^*\) framework [1]. The D-Case steps particularly focus on stakeholders and defines practical evaluation criteria for assurance cases written in GSN [7]. Our observation is as follows. To avoid argument divergence and difficulties of goal decomposition, we restrict the content of a GSN diagram only for particular stakeholders, and the way of goal decomposition is determined by preference of the stakeholders.

We are still in middle stages for facilitating use of assurance cases in Japan. However, several companies are now testing GSN and assurance cases in their work, and we have obtained several observations on the use of assurance cases.

The contributions of the paper are as follows.

- We define an assurance cases creation method called “D-Case Steps”, based on \(d^*\) [1] method.
- We open 8 GSN workshops using D-Case Steps, in which 113 engineers participated.
- We evaluate the effectiveness of D-Case Steps based on the results of workshops.
- We list findings and observations obtained from our a decade experience for facilitating use of assurance cases (more generally, facilitating use of software engineering methods).

The rest of the paper is organized as follows. In Sect. 2, we discuss related work. Section 3 introduces D-Case steps. Section 4 shows contents of GSN workshops and results of the workshops. Section 5 shows evaluation of the workshops. Section 6 summarizes lessons learned from our assurance cases facilitation work in Japan. Section 7 concludes the paper.

2. Related Work

The study of assurance cases have mostly varied in case studies such as [10], safety regulation [11], [12]. patterns [13], [14], confidence [15]–[17], and tool design and implementation [18]–[20]. A comprehensive survey of the study of assurance cases is shown in [21].

Despite of its significance, there have not so many guidance on creating assurance cases. Kelly proposed the six steps method for creating GSN (Goal Structuring Notation) diagrams [6]. Using the six step method, a GSN diagram is created in top-down (and also in bottom-up) manner. There are also several guidebooks on creating assurance cases, specially in GSN such as [3], [22], [23]. Such guidebooks are basically based on [6], and/or require knowledge on safety analysis and the system domain. We observe that the six steps method is an abstract method and needed to be elaborated for actual uses. Also, it is difficult for beginners to assume safety analysis and domain knowledge of the system for creating assurance cases.

In software engineering, it is common to consider stakeholders, their concerns, and relationships among them such as \(i^*\) framework [24] and KAOS [25]. However, in assurance cases, as far as we know, stakeholders (agents or actors) analysis have not been well discussed. One exceptional work of applying the notion of actors for assurance cases creation is \(d^*\) framework [1], [26]. In [1], [26], actors represent both systems and stakeholders involved in the scope. \(d^*\) framework represents two kinds of GSN: one kind is to show that an actor satisfies responsibility to other actors, and another kind is to show that an actor itself is dependable. The previous work [27] of the authors compared \(d^*\) framework and the methodology to construct assurance cases with the help of ArchiMate [28].

For practical use of assurance cases, there have been works on integrating assurance cases with existing methods. The work [29] uses GSN diagrams created in D-Case Editor [30] for hardware security evaluation. In [31], [32], GSN diagrams are used for compliance with Common Criteria in security. In [33], [34], GSN diagrams are used for explanation of the validity of safety analysis. The work [35] uses GSN diagrams for explication and evaluation of general safety requirements for space systems. The result indicates that GSN diagrams are effective for correcting errors in safety requirements.

Assurance cases are required for certification of safety-critical systems. If assurance cases are not linked with development and operational documents, assurance cases can not really be for “certification” of safety of the system. Otherwise, assurance cases are just for “paper safety” [11].
works introduced in this Section are mainly for integrating assurance cases with other risk analysis and software engineering methods, and the results of the works seem positive. However, unfortunately, assurance cases have not been widely used in Japan and other countries. Of course, safety cases, assurance cases for system safety are widely required for safety regulation in the United Kingdom, but the actual contents of assurance cases have not been widely publicized because they contain confidential information of safety-critical systems [36]. This makes assurance cases as a “black box”: we cannot see actual assurance cases and how they are created. Or even worse, assurance cases might not exist in the sense of the definition. The notion of assurance cases have been still getting attention as advanced systems such as artificial intelligence systems are becoming a crucial part for safety related system, and recent several works such as [37], [38] present argument structure for artificial intelligent systems. However, in such papers, only a high-level argument structure is presented, and actual and detail assurance cases for such systems have not been publicized: assurance cases are still black boxes for such systems.

This paper explores toward making assurance cases actual things: they can really be created by engineers and integrated with other risk analysis and documents by a method and workshops.

3. D-Case Steps

In this Section we introduce D-Case Steps. The novel features of the D-Case Steps are as follows.

- F1: Focusing on contexts.
- F2: Focusing on stakeholders.
- F3: Focusing on consensus building.

These features are derived from our experiences of creating and teaching assurance cases using GSN. System attributes such as dependability and safety are diverse, thus assurance cases can be huge and even unlimited. Also, choosing granularity of assurance cases are difficult. An assurance case can be just one goal “System is safe” with evidence “test results”, or hundreds of goals and evidence nodes through the level of reliability of one tiny resistance circuit.

For the problem of scale of assurance cases, first, we focus on contexts. Context information includes assumptions, premises, and system environment for the assurance argument of the target system. Without explicitly defining contexts, the argument becomes unlimited. In our D-Case Steps workshop, we focus on the necessity of context information as it determines the scope of an argument.

Next, we focus on stakeholders. Conventional assurance cases do not specify who creates an assurance cases and to whom the assurance case to be shown. However, as systems become huge and networked, it is difficult to fully understand the whole system by a single stakeholder. Therefore, we consider assurance cases in the relationship of stakeholders, and create an assurance case between a particular relationship of stakeholders such as the developer and the user of a system. For such a relation, the contents of the assurance cases can be focused only on the stakeholders concern, thus difficulty of creating the assurance case could decreased and the scale could be moderate. The whole assurance case of a system can be composed of the assurance case of each relationship of the stakeholders.

Third, we focus on consensus building among stakeholders. Conventionally, an assurance case is evaluated by several criteria such as completeness and confidence [39]. However, in such criteria, stakeholders are not considered. Assurance cases are for arguing dependability, safety, . . . , of systems among the stakeholders. Therefore it is important to evaluate that an assurance case can be really used for communication among the stakeholders, and make consensus on the dependability, safety, . . . , of the system.

D-Case Steps consist of three steps: Stakeholder analysis step, GSN writing step, and consensus building step. Figure 3 depicts the D-Case steps. Essentially, D-Case steps is a simplified version of d framework [1]. In [26], creating assurance cases of a system by d framework requires about 10 hours. We focus on creating assurance cases between particular stakeholders (in d framework, assurance cases between actors and of actors themselves are created), and avoid inter-relationship among assurance cases. This simplification enables us to practically create assurance cases within a few hours.

We use the following running example. This example is used in workshops for group work, and The example is called “Smart Room Viewing”†. The system is used for

†https://www.smartnairan.net/
viewing a room when a customer considers renting the room from the real estate company. The functions are as follows.

- Unattended viewing by advance reservation. Usually (in Japan) when renting a room, first the customer views the room accompanied with a real estate employee. However, using smart room viewing, the customer can view the room without one of the employees.
- Unlocking the room key is done by user authentication using customer’s smart-phone.
- Inside the room, the customer can check information of the room by a smart tablet.
- The customer can freely view the room within the time limits.
- The customer receives a notation of 5 minutes before the end by the tablet.
- The room is secured by surveillance cameras when customer is viewing the room.
- When finishing room viewing, the customer check that the room key is locked.

The flow of Smart Room Viewing system is shown in Fig. 4.

3.1 Stakeholder Analysis Step

To create an assurance case, first we need to analyze the stakeholders of the system (feature F2). An assurance case must be understandable to and concern of the stakeholders. Also, identifying stakeholders helps to limit the content of the GSN diagram. There are various methods for stakeholder analysis. Currently, D-Case steps do not specify which stakeholder analysis to use. However, at least the result of analysis should include the relationship between stakeholders. For example, a result of stakeholder analysis for Smart Room Viewing system is shown in Fig. 5.

From the stakeholder analysis, three stakeholders are identified: system provider, real estate company, and customer. The arrows among them indicate assurance relations. The system developer and real estate company have some claims which are needed to be assured to each other. For example, the system developer needs to assure security of Smart Room Viewing. Both the system developer and real estate company need to assure some claims to the customer.

3.2 GSN Creation Steps

After analyzing the stakeholders, GSN diagrams are created for the stakeholders. The creation steps consist of three sub steps, which are essentially the same steps of the six steps in [6], [7], except considering the stakeholders’ concerns.

1. Set the top goal and the contexts according to stakeholders’ concerns. For smart room viewing, there could be several top goal candidates:
   - The customer can obtain necessary information for renting the room. This claim is about information quality of the smart room viewing system.
   - Smart room viewing system is acceptably secure. This claim is assured in possibly three cases: by the system developer to the real estate company, by the system developer to the customer, and by the real estate company to the customer.
   - Smart room viewing system is cost-effective. This claim is assured by the system developer to the real-estate company.

Dependability of a system involves various aspects of the system: safety, security, information quality, cost-effectiveness, and so on. Each stakeholder has his or her own concerns and they are related to each other, and form the dependability of the system as a whole. D-Case steps specifies how to create a GSN diagram between particular stakeholders. Note that there could be multiple top goals.

2. Set the strategy from stakeholders’ interests and divide them into subgoals. This step is repeated until the stakeholders reach detailed evidence that can be accepted.

3. Set each evidence as the final leaf of the GSN.

In order to create a GSN diagram, it is necessary to set evaluation criteria of GSN. There are several criteria proposed in previous work. From such criteria, in D-Case step, we use the following three criteria.

- Context Validity (feature F1). Context nodes plays crucial rules in GSN diagrams, and contexts nodes should be linked to appropriate goals or strategies to describe context information of them. The importance of context is discussed in several previous work, such as [40] and [16], and we explicitly evaluate the validity of contexts.
- Logicality. The structure of a GSN diagram represents
a logical argument that the claim in the top goal holds. Therefore it is important to check the logicality of the GSN diagram.

- Relevance of Scale (feature F2). The stakeholders of a system have their own limited time and knowledge about the system. Thus it is important to limit the scale of the GSN diagram so that the stakeholders can read and understand it. As far as we know, this criteria has not been discussed in the literature.

There are other criteria proposed in the literature such as quantitative confidence [15], [17]. Based on our experience, we have set the above three evaluation criteria. In addition, it was adopted as criteria for ordinal average engineers to use easily.

3.3 Consensus Building Step (Feature F3)

After creating a GSN diagram, the stakeholders argue that the claim in the top goal is acceptable or not by checking the GSN diagram. From our experience, the number of stakeholders should be within five. Discussions will diverge when it comes to more stakeholders. In this step, the three evaluation criteria (context validity, logicality, and relevance of scale) are also used. Once all participated stakeholders make consensus that the claim is acceptable, then this step finishes. This step can be elaborated as shown in [7] or other methods such as Mind Map or FTA analysis.

3.4 Limitation of D-Case Steps

Current D-Case Steps are abstract and detailed procedures are not defined. In other words, we can elaborate D-Case Steps using conventional method. For example, Step 1 can be combined with other requirement analysis methods. Also, evaluation criteria for GSN diagram is tentative: in assurance case community, there have been still various argument for evaluation criteria.

4. GSN Workshops

Since 2012, we have opened assurance case workshops. The 8 workshops in 2018 used D-Case steps (Table 1). The goal of workshops is to make the participants to be interested in assurance cases and GSN, and they use assurance cases and GSN in their work. For the goal, we have continuously improved contents of workshops. The improvements are mainly on the basic structure of the workshops, contents of the group work, and associated facilitation techniques.

This paper reports the contents and evaluation of the 8 workshops.

In Table 1, “Open” means that the workshop was open call for participation. “Close” means that the participants are a closed group who are interested in system modeling such as UML. “V” is a company for third-party verification. “N” is a company for network systems. “T” is a company for information and communication systems.

Figure 6 shows a snapshot of the sixth workshop. The goals of workshops were as follows.

- (W0) The workshops themselves attract participants.
- (W1) Participants are interested in assurance cases, GSN, consensus building, and consider to apply GSN to their work.
- (W2) Participants can create appropriate GSN diagrams for a system or an agreement among different stakeholders by the features F1, F2, and, F3.

Rationale for W1 is as follows. We focus on consensus building (risk communication) among different stakeholders (feature F3), as it is a main role of assurance cases. Also, not only as certification documents, assurance cases should be used in practice to avoid paper safety [11].

The 8 GSN workshops mainly consist of the following contents.

1. Introduction of GSN and simple exercises for GSN syntax (40 minutes). We briefly introduce assurance cases and GSN syntax with simple exercises. We change the contents according to the interests of participants. For example, for workshop 2, we explain that GSN can be an argument model comparing with UML, SysUML, and other models. For workshop 4, 5, and 8, we explained that assurance case can be used for explaining how test results for systems contribute as evidence for dependability of the systems.

2. Introduction of D-Case steps (20 minutes).

3. Exercises for D-Case steps by group discussion (about 2 hours). Participants are divided into groups of 3 to 5
participants.

In the first 3 workshops, we had several trials and errors based on the participants’ feedback. For example, in these workshops, we used a web-based GSN editor called D-Case Communicator [19]. The tool works in conventional web browsers and enables users to remotely collaborate on creating a GSN diagram. Specially, by using the tool, a GSN diagram created by a participant can be immediately shown in the screen of the projector. However, we found that participants need some time to get used to the tool, and the time spent tends to stop conversations in the group. Also, by using a tool, only one or two participants can create GSN diagrams, and other participants cannot join in the creation. Based on the observation, for the next 5 workshops, we stop using the tool. Furthermore, during the first 3 workshops, we recognize the need for facilitation. Participants of a workshop are not necessarily known to each other. Therefore, facilitation techniques such as ice break and self-introduction are important for the success of workshops.

4.1 Workshop Contents

For the success of workshops, choosing contents of exercises is important. We mainly select the following 4 contents according to closeness to the practice in the software development companies.

- “Planning for travelling” is the most unfamiliar content for software engineers, but it is a familiar situation if the participants spent college life in their youth.
- “Buying home appliances” in a home of husband and wife with both working with a baby is also a familiar situation for engineers.
- “Smart Room Viewing System” is a relatively new system, but most of software engineers could understand the specification.
- “Testing for Car Navigation System” is the most practical contents which is based on actual case study in a test and verification company.

Based on the feedbacks from previous workshops, we change contents of the next workshop. The contents of workshops 1 through 8 is shown in Table 2. In this paper we omit the detail of the three contents of workshop 1.

The rationale for these 4 contents is as follows. First, workshops should attract the participants. Second, the contents should be understandable to the participants, whose backgrounds and works vary, even in the same company. Therefore, if content of a workshop is too specific, most of the participants (except for the expert) would lose interest to the group work. Or, if the contents of group work is too general from participants work, then the participants would also lose interests for the workshop.

4.2 Workshop Results

For each workshop, we collect results of questionnaires, GSN diagrams created by the participants (except for workshop 1, as we could not collect answer sheets of the participants), and comments from the participants. During the workshops, we simultaneously developed the contents and questionnaires items to the participants, and the questionnaires items were not same for all the workshops. In addition, workshops 4, 5, 6, and 8 are held as the company’s employee training, we basically use the company’s own questionnaires items. Also, we add ARCS model [41] questionnaires. ARCS (“Attention, Relevance, Confidence, and Satisfaction”) model as it is an instruction model for motivating for learning, and it is often used for evaluation of workshops.

4.2.1 Questionnaires

In summary, we obtain the following results of questionnaires.

- For workshop 1: contents of workshop (Q1) and practicability (Q2).
- For workshop 2 and 3: contents of workshop (Q1), practicability (Q2), understandability (Q3), easiness of creating top goal (Q4), context (Q5), strategies (Q6), goal decomposition (Q7), evidence (Q8), and applicability to consensus building (Q9).
- For 4, 5, 6, 7 and 8: ARCS (attention (QA), relevance (QR), confidence (QC), satisfaction (QS)), practicability (Q2), understandability (Q3), and recommendation to others (Q10).

The results are as follows. Each result shows the average of answers by the participants which are normalized in [0, 1].

- Workshop 1 (Table 3).
- Workshop 2 and 3 (Table 4).
- Workshop 4, 5, 6, 7, and 8 (Table 5)

Figure 7 shows a graph of the ARCS models results in

| Table 2 | Contents of D-case workshops in FY2018 |
|---------|---------------------------------------|
| 1       | Robot car contest, IoT system, and test pattern |
| 2       | Smart Room Viewing system |
| 3       | Smart Room Viewing system |
| 4       | Buying home appliance and Testing for Car Navigation System |
| 5       | Planning for traveling |
| 6       | Smart Room Viewing system |
| 7       | Smart Room Viewing system |
| 8       | Smart Room Viewing system |

| Table 3 | Questionnaires results of workshop 1 |
|---------|-----------------------------------|
| Q1      | Q2 |
| 0.84    | 0.82 |

| Workshop 2 | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 |
|------------|----|----|----|----|----|----|----|----|----|
| 0.88       | 0.78 | 0.78 | 0.80 | 0.65 | 0.68 | 0.70 | 0.73 | 0.80 |

| Workshop 3 | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 |
|------------|----|----|----|----|----|----|----|----|----|
| 0.88       | 0.85 | 0.78 | 0.80 | 0.65 | 0.68 | 0.70 | 0.73 | 0.80 |
Table 5 Questionnaires results of workshop 4, 5, 6, 7, and 8

| Workshop | QA   | QR   | QC   | QS   | Q2   | Q3   | Q10  |
|----------|------|------|------|------|------|------|------|
| Workshop 4 | 0.61 | 0.61 | 0.48 | 0.63 | 0.67 | 0.69 | 0.57 |
| Workshop 5 | 0.57 | N/A  | 0.54 | 0.69 | 0.68 | 0.57 | 0.57 |
| Workshop 6 | 0.94 | 0.87 | 0.70 | 0.96 | 0.94 | 0.88 | 0.92 |
| Workshop 7 | 0.95 | 0.95 | 0.64 | 0.98 | 0.95 | 0.79 | 0.86 |
| Workshop 8 | 0.92 | 0.96 | 0.67 | 0.96 | 0.90 | 0.69 | 0.90 |

Table 5 for the workshops 4, 5, 6, 7, and 8. Note that for QR of Workshop 5, we could not collect the date due to our mistake.

4.2.2 Comments

For each workshop, we asked the participants to write free comments about the workshop. We categorized the collected comments by relating them to workshop goals W0, W1, and W2. For example, a comment “the workshop was interesting” is categorized into W0. Next, we label each comment positive or negative. For example, a comment “GSN can be used for consensus building” in W1 is labeled positive and a comment “I can not find any application of GSN in my work” in W1 is labeled negative. Table 6 shows the results of categorization and labeling of each workshop.

In Table 6, numbers of positive and negative comments are shown as pairs. For example, the pair (1, 4) for W0 of Workshop 1 means 1 positive and 4 negative comments.

4.2.3 GSN Diagrams

From the 8 workshops, we obtained 33 GSN diagrams created by the participants as follows. We denote “smart room viewing system” as content 1, “buying home appliance” as content 2, “testing for Car navigation system” as content 3, and “planning for traveling” as content 4, respectively.

- Workshop 2: GSN diagrams 1, 2, 3, 4, 5 for content 1.
- Workshop 3: GSN diagrams 6, 7, 8, 9, 10 for content 1.
- Workshop 4: GSN diagrams 11, 13, 15, 17, 19, 21 for content 2 and GSN diagrams 12, 14, 16, 18, 20, 22 for content 3.
- Workshop 5: GSN diagrams 23, 24, 25 for content 4.
- Workshop 6: GSN diagrams 26, 27, 28, 29, 30 for content 1.
- Workshop 7: GSN diagram 31 for content 1.
- Workshop 8: GSN diagrams 32 and 33 for content 1.

The averages of number of nodes for the workshops are shown in Fig. 8.

We asked 5 engineers and researchers (A, B, C, D, and E) who have some experiences on assurance cases and GSN (3 to 10 years) to evaluate the 33 GSN diagrams. The evaluation criteria are context validity and logicality of argument structure, which are evaluation criteria of D-Case steps. Both evaluation are done by 5 level rating (1 is lower and 5 is higher). Figure 9 shows the result of evaluation, which shows the total of context validity and logicality of argument. The horizontal and vertical axes are number of GSN diagrams and total score of the GSN diagrams, respectively.

The average of standard deviations of scores of the evaluators for the 33 GSN diagrams is 1.2.

Figure 10 shows the evaluation results by averages of GSN diagrams of each workshop and evaluator. The horizontal and vertical axes are number of the workshops and average of scores (context validity, logicality of argument, and the total) of the GSN diagrams, respectively.
5. Evaluation

5.1 Questionnaires

The questionnaires are classified for each workshop goals W0, W1, and W2 as follows.

- W0: Q1 and QS.
- W1: Q2, Q9, Q10, QA, and QR.
- W2: Q3, Q4, Q5, Q6, Q7, Q8, QC.

5.1.1 Evaluation of W0

The points of Q1 and QS are mostly high (0.84 to 0.98) except for QS of workshops 4 and 5. This indicates that the workshops 4 and 5 had some problems.

5.1.2 Evaluation of W1

The points of Q2, Q10, QA, and QR are mostly high (0.78 to 0.96) except for those of workshops 4 and 5. Q9 is only asked to the participants of the workshop 2 and 3, but the points (0.8 for both workshop 2 and 3) indicate that the participants are positive on using the method for consensus building.

5.1.3 Evaluation of W2

For those questionnaires, the points are relatively low comparing with other questionnaires (the highest points are 0.80 for Q4 of the workshop 2 and 3). In particular, QC for the workshop 4 and 5 are very low (0.48 and 0.54, respectively).

5.2 Comments

The positive comments related to W0 include such that the workshop is interactive and group discussion is fun. The negative comments related W0 are such that the workshop is not well organized. In particular, for workshop 4, several participants wrote that the time for exercise were not enough (two hours for two exercises). Also, some participants of workshop 4 and 5 complained that the contents (buying home appliance and planning for travelling) were not appropriate because they are not related to their work.

The comments related to W1 includes several possible application of GSN and the D-Case steps as follows.

- System assurance.
- Making explanation documents for customer, manager, etc.
- Personal use: use with logic tree, logical structuring ideas.
- Consensus Building among different sectors.
- Product planning. Using with risk analysis in project starting, and Using in very early stage of requirement analysis

System assurance is the original purpose of assurance cases. These are positive results that the participants consider other usages. On the other hand, there were negative comments such that the participants could not find any application of GSN and the method. Such negative comments were mostly from workshop 4.

The comments related to W2 were mostly about the difficulties of creating GSN diagrams using goal, strategy, evidence, and context nodes. Among them, 25 comments were about difficulties in context nodes, 19 in strategy nodes, and 14 in goal nodes. The results indicate that creating context nodes is most difficult.

5.3 GSN Diagrams

The GSN diagrams created by the participants are used to evaluate workshop goal W2. The average standard deviation of evaluation of 5 evaluators is 1.2, which is relatively low. This indicates that although each evaluator has his own evaluation method, it shows possibility of developing a fair evaluation method using context validity and logicality of argument.

As in Fig. 9 and Fig. 10, the results of the workshop 4 and 5 are relatively low to that of the other workshops. For other workshops using “smart room viewing system” (workshops 2, 3, 6, 7, and 8), the averages of scores are increasing (except for workshop 8, in which one of the two groups
failed to create a GSN diagram). This indicate our workshops have been improved by the contents, lecture method, and facilitation.

The average of number of GSN nodes by the participants (Fig. 8) indicates that at most a GSN diagram with 20 nodes can be created by engineers in a working time (2 hours). It is clear that workshop 4 (two exercises in two hours) had not enough time for exercises.

5.4 Features of D-Case Steps

The features F1 and F2 of D-Case steps are for limiting contents of contexts and goals and creating assurance cases. In Fig. 9 and Fig. 10, some participants could create GSN diagrams which got high scores from the evaluators who have 3 to 10 years experiences on assurance cases. This is by (at least partly) F1 and F2. For example, we show the GSN diagram GSN 31 which got the highest score 8.4/10.0 in Fig. 11. GSN 31 diagram clearly considers the stakeholders (the company and the user) and argues how “smart” is the smart room viewing system (F2) by decomposing the top goal by each smartness of the system, by stating appropriate contexts (F1). The comments of the evaluators for GSN 31 are as follows: “amount of information is appropriate and the argument structure is convincing,” “the basic of GSN diagram is sufficient,” and “the context nodes eliminate ambiguity and the argument is clear.”

The feature F3 is briefly evaluated by Q9 (the results were 0.8 in both workshop 2 and 3) and a few positive comments from the participants. To evaluate F3 in detail, we consider that the workshops are needed to include role playing exercises between stakeholders (such as a developer and a customer) for making consensus building (such as “a system is dependable”) using the created GSN diagrams.

5.5 Summary of Evaluation

Overall, the workshop 4 and 5 seems to had problems on the structure of workshop as two hours for two exercises were not enough. Also, as the contents are not directly related to their work (in particular, “Planning for Traveling” was far from their expectation), it was difficult for the participants to consider contexts and goals. For the exercise “smart room viewing system,” there were no negative feedbacks from the participants. The content is also not directly related to the work of the participants, but it seems to be not difficult for them to consider the smart room viewing system. The number of participants of workshop 4 (28 participants) was also problematic, as the lecturers were difficult to grasp activities of all groups. The structure of workshops affects all of W0, W1, and W2.

From the evaluation of W2, the difficulty of creating assurance cases still remains. By improving contents of workshops, however, from Fig. 9 and Fig. 10 in which scores are increasing for “smart room viewing system” exercise, it is indicated that the difficulty can be reduced by improving content of D-Case workshop.

5.6 Limitation

A limitation of above evaluation is that backgrounds of participants are diverse, and conditions for subject experiments
such as equality of participants knowledge and experiences are not assumed. Indeed, there are many parameters such as contents of workshops and time spent for group discussions. Instead, our results are obtained from working engineers who are interested in assurance cases and considering to apply assurance cases to their work. Therefore, comments and answers for questionnaires from them can be said as more practical (but not precise as usual subject experiments) evaluation of our method.

6. Lessons Learned from Assurance Cases Facilitation Activities

The workshops held in 2018 and 2019 were based on the observations obtained from our a decade experiences for facilitating use of assurance cases in Japan. At first, we thought that as assurance cases has been widely used in the UK, it would be easy to introduce them in Japan, as we can use existing methods and tools developed in the UK. Unfortunately, however, it was not: the content of assurance cases were strictly confidential[36], and there were only a few publically available assurance cases in the Internet; only abstract methods were available[7]; also, only English documents were available at that time. Therefore we needed to begin our activities from creating Japanese documents almost from scratch.

We started our work in JST CREST DEOS project. The DEOS (Dependable Embedded Operating System) project, which is started in 2006 and ended in 2012 [42], and has been continuing as a consortium, was a Japanese national project for developing dependable systems. The project aims to use assurance cases in both the development and operational phases (specially for failure response action) ([43], [44]) for assuring the dependability of the target systems. At that time, the problem of complexity of systems and diversity of stakeholders were already recognized by Japanese industries. When we introduced assurance cases to them, they believed that assurance cases are a promising approach for solving the problem. Since then we did research and practical realization of assurance cases in Japan. Our activities can be modeled as in Fig. 12.

First, we do research on a software engineering method. Next, to facilitate use of the method in companies, we open open workshops. For this we prepare contents (tutorial, exercises, and workshop facilitation) and websites, etc. Third, if a participant from a company gets interested in our method, we discuss with the participant about how to use the method in the company. If preliminary use of the method in the company is positive, we open closed workshops inside the company. If the results are positive, the company decides to use the method regularly. During these steps, we continuously get feed-backs from participants and companies and update the method. Following the model, we had six close workshops in four companies. Our method has not reached to the final step, but several companies have been conducting preliminary use of our method and closed workshops with us toward the final step:

- An automotive company is preliminary using our method for communicating with a foreign alliance company and in participating standardization activity of an automotive safety standard.
- Fuji Xerox uses our method in ET robot contest†, and won the championship.
- A group of JAXA uses GSN for communicating with space development companies. We had closed workshop inside JAXA in 2013, and have been continuing discussion on GSN†† with them.
- Company V had three GSN workshops in 2019 by themselves based on our workshops.
- Company N is trying to introduce assurance cases, and as a first step, they are preparing to introduce GSN tool called Astah GSN††† in the company.
- Company T is trying to use our method in very early stage of requirement analysis.
- Mitsubishi Electric Engineering is testing GSN††††.

From our experiences of communicating with a few hundreds of engineers, introducing a software engineering method which involves stakeholder and analysis to a company is difficult. One company shared the following necessary conditions with the authors for successfully introducing assurance cases to industry engineering processes.

- To obtain an understanding on necessity of accountability
- To share cost-benefit performance on assurance cases between the management layers and development teams
- To provide any education and training programs for skills on utilizing assurance cases to engineers
- To provide any measures (e.g. software tools) for utilizing assurance cases to engineering teams

We list our observations for the reasons.

- Engineers do not change their accustomed way of work and do not want to use a new method. They do not have time to study new methods.
- Managers do not want to introduce a new method to the developing environment unless the method is proven to be effective by quantitative evaluation. For example, “the method reduces the development cost by 20%” must be proven.
- In a company, most of the engineers work in a closed environment. Therefore, they do not need to consider

†https://www.etrobo.jp
††www.dcase.jp
†††http://astah.net/release-notes/gsn/gsn-1.2
††††https://ja.astahblog.com/2019/05/08/case_study_for_gsn/
stakeholders outside of the section or the company, and difficult to recognize the need of identifying diverse stakeholders.

- Explicit logical analysis such as goal oriented methods is not familiar to engineers. In a goal decomposition in GSN, even the simple condition that if all sub goals are satisfied, then the goal is satisfied seems not easy for some engineers, as discussed in [8].

To introduce a new method to a company, we need to solve these difficulties. In particular, we must convince engineers and managers that the method is really helpful to solve their issues from their point of view, not from researchers view. Also, workshop construction and facilitation are important, as workshops are the first opportunity for getting to know the method. One of findings is that number of participants of a workshop should be within 20. Otherwise, we can not control the workshop as participants discussions diverge.

Recently, as systems become complex and huge and various stakeholders are involved, and new kinds of systems such as artificial intelligence systems become to be used, notions of software engineering such as system assurance, stakeholders, goal oriented analysis, ..., become more important. We believe that assurance cases and software engineering methods should be used more widely.

7. Concluding Remarks

This paper has presented D-Case steps, a new method for creating assurance cases. D-Case steps focuses on stakeholders, and includes conventional GSN creating steps and consensus building steps. We had GSN workshops held in Japan, and the feedbacks from the participants have been becoming positive. We are planning to use the D-Case steps for writing GSN in conventional use such as of functional safety of automobiles, specially autonomous automobiles, for which various stakeholders are related. We would like to continue our activities on assurance cases, and present our progress near future.

Acknowledgments

We would like to thank students in our laboratories for cooperating workshops with us. In particular, Yuto Onuma mainly contributed to development of the contents of the workshops and collected data used in the paper. Also, we have significant assistance from industrial collaborators: Ichiro Yamaura, Hideki Nojo, Koichi Akiyama, Isao Saito, and Tsutomu Koshiyama, to name a few. Last but not least we express sincere gratitude to people of the DEOS consortium including Yoshiki Kinoshita, Makoto Takeyama, and Mario Tokoro.

References

[1] S. Yamamoto, Y. Matsuno, and M. Tokoro, “d* framework: Inter-depending model for dependability (fast abstract),” IEEE/IFIP International Conference on Dependable Systems and Networks, DSN 2012, Boston, MA, USA, June 25-28, 2012, pp.1–2, 2012.
[2] Y. Onuma, T. Takai, T. Koshiyama, and Y. Matsuno, “D-case steps: New steps for writing assurance cases,” Proc. Computer Safety, Reliability, and Security - SAFECOMP 2018 Workshops, ASSURE, DECSoS, SASSUR, STRIVE, and WAISE, Västerås, Sweden, Sept. 18, 2018, pp.71–78, 2018.
[3] European Organisation for the Safety of Air Navigation, “Safety case development manual.” European Air Traffic Management, 2006.
[4] Railtrack, “Yellow book 3,” Engineering Safety Management Issue3, vol.1, vol.2, 2000.
[5] R. Bloomfield and P. Bishop, “Safety and assurance cases: Past, present and possible future - an Adelard perspective.” Proc. Eightheenth Safety-Critical Systems Symposium, Bristol, UK, 2010.
[6] T. Kelly, “Arguing Safety - A Systematic Approach to Safety Case Management,” Ph.D. thesis, Department of Computer Science, University of York, 1998.
[7] Assurance Case Working Group, “Goal structuring notation community standard version 2,” Jan. 2018. downloaded from https://ascn.uk/ir141B-1.
[8] S. Yamamoto and Y. Matsuno, “An evaluation of argument patterns to reduce pitfalls of applying assurance case,” Proc. 1st International Workshop on Assurance Cases for Software-Intensive Systems, ASSURE ’13, San Francisco, California, USA, May 19, 2013, pp.12–17, 2013.
[9] Y. Matsuno, J. Nakazawa, M. Takeyama, M. Sugaya, and Y. Ishikawa, “Toward a language for communication among stakeholders,” Proc. 16th IEEE Pacific Rim International Symposium on Dependable Computing (PRDC’10), pp.93–100, 2010.
[10] R. Hawkins, K. Clegg, R. Alexander, and T. Kelly, “Using a software safety argument pattern catalogue: Two case studies,” Proc. Computer Safety, Reliability, and Security - 30th International Conference, SAFECOMP 2011, Naples, Italy, Sept. 19-22, 2011, pp.185–198, 2011.
[11] N. Leveson, “The use of safety cases in certification and regulation,” ESD Working Paper Series, Boston: MIT, 2011.
[12] J. Birch, R. Rivett, I. Habli, B. Bradshaw, J. Botham, D. Higham, P. Jesty, H. Monkhouse, and R. Palin, “Safety cases and their role in ISO 26262 functional safety assessment,” Proc. Computer Safety, Reliability, and Security - 32nd International Conference, SAFECOMP 2013, Toulouse, France, Sept. 24-27, 2013, pp.154–165, 2013.
[13] R. Alexander, T. Kelly, Z. Kurd, and J. McDermid, “Safety cases for advanced control software: Safety case patterns,” tech. rep., Department of Computer Science, University of York, 2007.
[14] T. Kelly and J. McDermid, “Safety case patterns - reusing successful arguments,” IEE Colloquium on Understanding Patterns and Their Application to System Engineering, 1998.
[15] J.B. Goodenough, C.B. Weinstock, and A.Z. Klein., “Toward a theory of assurance case confidence,” tech. rep., Carnegie Mellon, 2012.
[16] P.J. Graydon, “Towards a clearer understanding of context and its role in assurance argument confidence,” Proc. Computer Safety, Reliability, and Security - 33rd International Conference, SAFECOMP 2014, Florence, Italy, Sept. 10-12, 2014, pp.139–154, 2014.
[17] L. Duan, S. Rayadurgam, M.P.E. Heimdahl, O. Sokolsky, and I. Lee, “Representation of confidence in assurance cases using the beta distribution,” 17th IEEE International Symposium on High Assurance Systems Engineering, HASE 2016, Orlando, FL, USA, Jan. 7–9, 2016, pp.86–93, 2016.
[18] E. Denney, G. Pai, and J. Pohl, “Advocate: An assurance case automation toolset,” Proc. Computer Safety, Reliability, and Security - SAFECOMP 2012 Workshops: Sassur, ASCoMS, DESEC4LCCI, ERCIM/EWICS, IWDE, Magdeburg, Germany, Sept. 25-28, 2012, pp.8–21, 2012.
[19] Y. Matsuno, “D-case communicator: A web based GSN editor for multiple stakeholders,” Proc. Computer Safety, Reliability,
and Security - SAFECOMP 2017 Workshops, ASSURE, DECSoS, SASSUR, TELERISE, and TIPS, Trento, Italy, Sept. 12, 2017, pp.64–69, 2017.

[20] Y. Matsuno, “A design and implementation of an assurance case language,” 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, DSN 2014, Atlanta, GA, USA, June 23-26, 2014, pp.630–641, 2014.

[21] J. Rushby, “The interpretation and evaluation of assurance cases,” Tech. Rep. SRI-CSL-15-01, Computer Science Laboratory, SRI International, Menlo Park, CA, July 2015, Available at http://www.csl.sri.com/users/rushby/papers/sri-csl-15-1-assurance-cases.pdf.

[22] C.B. Weinstock, “Assurance cases,” Dec. 2008. downloaded from http://www.seas.upenn.edu/~lee/09cis480/lec-AssuranceCasesTutorial.pdf.

[23] Adelard, Adelard Safety Case Development Manual, Adelard, 1998.

[24] E. Yu, “Social modeling and †,” Conceptual Modeling: Foundations and Applications - Essays in Honor of John Mylopoulos, pp.99–121, 2009.

[25] A. Dardenne, A. van Lamsweerde, and S. Fickas, “Goal-directed requirements acquisition,” Sci. Comput. Program., vol.20, no.1-2, pp.3–50, 1993.

[26] T. Saruwatari, “A Study on Assurance Cases Construction based on the Notion of Actors,” (in Japanese), Ph.D. thesis, Institute of Information Security, 2013.

[27] Q. Zhi, Z. Zhou, and S. Yamamoto, “Visualized assurance approach for enterprise architecture,” J. Inf. Commun. Converg. Eng., vol.17, no.2, pp.117–127, June 2019.

[28] The Open Group, “Archimate 3.0 specification,” 2016.

[29] H. Kawakami, R. Gallo, R. Dahab, and E. Nascimento, “Hardware security evaluation using assurance case models,” 10th International Conference on Availability, Reliability and Security, ARES 2015, Toulouse, France, Aug. 24-27, 2015, pp.193–198, 2015.

[30] Y. Matsuno, H. Takamura, and Y. Ishikawa, “A dependability case editor with pattern library,” Procs. IEEE 12th International Symposium on High-Assurance Systems Engineering (HASE), pp.170–171, 2010.

[31] T. Kaneko, “CC-Case: An Integration Method of Security Requirement Analysis and Assurance,” (in Japanese), Ph.D. thesis, Institute of Information Security, 2014.

[32] S. Yamamoto, T. Kaneko, and H. Tanaka, “A proposal on security case based on common criteria,” Proc. Information and Communication Technology - International Conference, ICT-EurAsia 2013, Yogyakarta, Indonesia, March 25-29, 2013, pp.331–336, 2013.

[33] N. Kobayashi, “A Study on Dependability Assurance Methods for In-Vehicle Systems,” (in Japanese), Ph.D. thesis, Nagoya University, 2018.

[34] N. Kobayashi, S. Morisaki, and S. Yamamoto, “The method of explanation for safety analysis results using d-case,” Journal of Information Processing Society of Japan, vol.58, no.2, pp.521–530, 2017.

[35] K. Kakimoto, S. Kawaguchi, T. Takai, N. Ishihama, H. Iida, and M. Katahira, “Explication and evaluation of general safety requirements using gsn,” SEC Journal, vol.12, no.3, pp.16–23, 2016.

[36] R. Steinzor, “Lessons from the North Sea: Should “Safety Cases Come to America?,” B.C. Envtl. Aff. L. Rev., vol.38, pp.417–444, 2011.

[37] L. Gaurerhuf, P. Munk, and S. Burton, “Structuring validation targets of a machine learning function applied to automated driving,” Proc. Computer Safety, Reliability, and Security - 37th International Conference, SAFECOMP 2018, Västerås, Sweden, Sept. 19–21, 2018, pp.45–58, 2018.

[38] F. Ishikawa and Y. Matsuno, “Continuous argument engineering: Tackling uncertainty in machine learning based systems,” Proc. Computer Safety, Reliability, and Security - SAFECOMP 2018 Workshops, ASSURE, DECSoS, SASSUR, STRIVE, and WAISE, Västerås, Sweden, Sept. 18, 2018, pp.14–21, 2018.

[39] R.E. Bloomfield, B. Littlewood, and D. Wright, “Confidence: Its role in dependability cases for risk assessment,” DSN, pp.338–346, 2007.

[40] J. Spriggs, GSN - The Goal Structuring Notation: A Structured Approach to Presenting Arguments, Springer, 2012.

[41] J. Keller, Motivational Design of Instruction, Hillsdale, NJ: Lawrence Erlbaum Associates, 1983.

[42] M. Tokoro, ed., Open Systems Dependability: Dependability Engineering for Ever-Changing Systems, CRC Press, 2012.

[43] Y. Matsuno and S. Yamamoto, “An implementation of GSN community standard,” Proc. 1st International Workshop on Assurance Cases for Software-intensive Systems (ASSURE 2013), pp.24–28, 2013.

[44] H. Fujita, Y. Matsuno, T. Hanawa, M. Sato, S. Kato, and Y. Ishikawa, “DS-Bench toolset: Tools for dependability benchmarking with simulation and assurance,” IEEE DSN, 2012. 9 pages.