

Algebraic Approach for Confidence Evaluation of Assurance Cases

Yoriyuki Yamagata¹ and Yutaka Matsuno²

¹ National Institute of Advanced Industrial Science and Technology (AIST),
1-8-31 Midorigaoka, Ikeda, Osaka 563-8577, Japan
yoriyuki.yamagata@aist.go.jp

² College of Science and Technology, Nihon University,
274-8501 Narashinodai, Funabashi, Chiba 7-24-1, Japan
matsuno.yutaka@nihon-u.ac.jp

Abstract. This paper presents a preliminary study on a method to evaluate the confidence of assurance cases using an abstract algebra mapped to a partial order. Unlike conventional quantitative methods for confidence evaluation, our approach is purely qualitative and employs a small number of axioms. It does not rely on numerical parameters that are difficult to determine in practice. Furthermore, our method can be regarded as an abstraction over numerical methods that use probability. To illustrate that our method provides a rigorous foundation for the qualitative evaluation of assurance cases, we give a sufficient condition for a multi-legged argument to improve confidence. Finally, we use our method to evaluate a concrete goal structuring notation (GSN) diagram that argues that a computer simulation of a biological system is reliable. These findings suggest that methods based on abstract axioms are viable approaches for confidence evaluation of assurance cases.

Keywords: Assurance case · Goal structuring notation (GSN) · Confidence · Formal semantics

1 Introduction

Creating and evaluating assurance cases are challenging tasks. The concept of assurance cases is given an abstract definition such as “A reasoned and compelling argument, supported by a body of evidence, that a system, service or organization will operate as intended for a defined application in a defined environment” [2]. By such an abstract definition, virtually all artifacts of the target system could be parts of the assurance case. In an automotive assurance case guideline [10], it is noted that “the question of when to stop adding further detail to an assurance case is not one that can be easily answered.” An important issue is understanding how much *confidence* one can have in the claim and how different arguments contribute to such confidence, given a claim and a supporting argument [6].

Goal structuring notation (GSN) [2] is a widely used graphical notation for assurance cases. A GSN diagram starts with a top goal of a claim to be argued,

such as “System X is safe.” Each goal is decomposed into sub-goals via a strategy node, which explains why the sub-goals are sufficient to support the goal, and finally, into directly verifiable evidence. A GSN diagram also documents the assumptions and contexts for an assurance case. This study uses GSN diagrams for presenting assurance cases.

Previous studies on the confidence of assurance cases were mostly based on numerical evaluations. A drawback of numerical evaluations is that the results depend on the numerical parameters used in the evaluation, whose appropriateness is difficult to verify. Thus, a widely applicable quantitative method for the confidence evaluation of assurance cases remains to be established.

This paper proposes a method to evaluate the confidence of assurance cases using an abstract algebra mapped to a partial order. This method has several advantages over numerical methods. First, the proposed method is defined using a small number of axioms without numerical parameters; thus, the results have a clear meaning. Second, the proposed method is based on weaker assumptions and is more general than numerical methods. Finally, the proposed framework can be regarded as an abstraction over (previously proposed) probability methods of confidence evaluation. Although our axioms are still weak for fine confidence evaluation, we believe that the method based on abstract axioms is shown to be a viable research direction.

The remainder of this paper is organized as follows. Section 2 reviews related studies. Section 3 gives the definition of GSN diagrams. Section 4 defines an abstract algebra of “states” and introduces our evaluation method. Section 5 relates our method to probabilistic evaluation. Section 6 describes the application of our method to *multi-legged arguments* [3] and gives a sufficient condition for a multi-legged argument to improve confidence. Section 7 analyzes a concrete GSN diagram that argues the correctness of a computer simulation of a biological process, namely lymphoid tissue formation. Finally, Section 8 states the conclusions and explores possible extensions of our framework.

2 Related Work

Developing an evaluation method for an assurance case is a current research objective, and various approaches have been proposed for this purpose. Studies on evaluating assurance cases are mainly concerned with the term *confidence*, i.e., how stakeholders gain sufficient confidence for the dependability of the system from the assurance cases. In [4], probability was used to calculate the confidence in assurance cases. Using the probabilistic framework, Bloomfield et al. [3] showed that independent multi-legged arguments increase the confidence of assurance cases. Other approaches include using Baconian probability [14] and the Dempster-Shafer theory (DST) [13]. The Baconian approach considers the doubts eliminated by the claims and evidence in a case with eliminative induction. DST supports the assignment of weights (i.e., mass) to a combination of possible events, rather than only assigning weights (probabilities) to each event, as is done in standard probability theory. In [9], Rushby noted that the

interpretation of assurance cases could be “inductive,” i.e., the conjunction of sub-claims strongly suggests that the claim is true, or it could be “deductive,” i.e., the conjunction implies (or entails or proves) the claim, and he emphasized that for evaluating assurance cases, the inductive nature of assurance cases must be considered.

The method proposed in this paper is based on an abstract algebra and a partial order; thus, it is purely qualitative. We do not regard the qualitative nature of our method as a drawback, because the reasoning of assurance cases is inherently qualitative: assurance cases are written in natural language (even using graphical notations) and thus exhibit a qualitative nature. Our method provides a rigorous theoretical foundation for the qualitative evaluation of assurance cases. Furthermore, it can be regarded as an abstraction of other methods. In Section 5, we show that all evaluation methods based on probability are special cases of our method if they obey the axioms of probability and certain weak conditions.

3 Goal Structuring Notation (GSN)

GSN is a graphical notation for representing an informal argument. The argument is constructed in a top-down manner, in which a “goal,” i.e., a final claim, is gradually elaborated as sub-goals and directly verifiable evidence.

A GSN diagram is constructed with the following types of nodes. A goal is

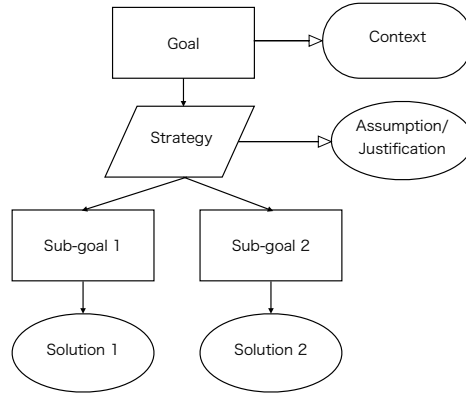


Fig. 1. GSN components

a claim to be demonstrated. A strategy is a method for deriving a goal, which decomposes the goal into several sub-goals (premises). A (sub-)goal may be demonstrated by direct evidence (solution). A context indicates an environment

that specifies how a goal is interpreted. An assumption and a justification are underlying reasons why a strategy is correct and taken for granted. We use the notation \diamond , which indicates an “undeveloped” argument, i.e., an argument that is not completed.

To facilitate the formal analysis, we introduce a term notation for GSN diagrams. In the definition, we omit the context and assumption nodes for simplicity of presentation.

Definition 1 (Term notations for GSN diagrams D , modified from [8]).

$$D ::= \langle g, \diamond \rangle \tag{1}$$

$$| \langle g, e \rangle \tag{2}$$

$$| \langle g, \text{OR}, (D_1, \dots, D_n) \rangle \tag{3}$$

$$| \langle g, \text{st}, (D_1, \dots, D_n) \rangle \tag{4}$$

$\langle g, \diamond \rangle$ is an undeveloped argument, and $\langle g, e \rangle$ is an argument directly derived from an evidence e . $\langle g, \text{OR}, (D_1, \dots, D_n) \rangle$ is a *multi-legged* argument, in which the same conclusion is obtained using different diagrams D_1, \dots, D_n . We discuss multi-legged arguments in Section 6. $\langle g, \text{st}, (D_1, \dots, D_n) \rangle$ is an argument derived from D_1, \dots, D_n using a strategy st .

4 Truthmaker semantics and confidence

4.1 Truthmakers

Classical logic assumes the principle of bivalence: the statement either holds or not, and there is no middle ground. However, the goals of GSN may not be interpreted to have just two truth values because we may not have sufficient information to determine the truth values. To interpret the goals of GSN, we adopt a radically different approach called *truthmaker semantics* [5], which has recently been developed in the field of logic.

Truthmaker semantics assumes that the world consists of objects called *truthmakers* that make a statement true. For example, the truthmaker of the statement “New York is rainy” is the rain in New York.

Truthmakers have a mereological structure [11], which represents a part-whole relation between two truthmakers. For example, the rain in New York is a part of the rain and wind in New York. The part-whole relation between the truthmakers forms an order relation. Further, we may amalgamate two truthmakers, say, the rain in New York and the wind in New York. The amalgamation of two truthmakers is represented by a binary operation \oplus .

Depending on their mereological structure, truthmakers obey different sets of axioms. In this study, we employ a small set of axioms to interpret a wide variety of confidence evaluations. Let \mathcal{S} (a state-space) be a set of truthmakers that we consider.

Definition 2. *Let \mathcal{S} be a state space. Then, \mathcal{S} has the element 0, the binary operator \oplus , and the order relation \sqsubseteq such that*

1. (Unit) $0 \oplus s = s \oplus 0 = s$.
2. (Commutativity) $s_1 \oplus s_2 = s_2 \oplus s_1$, $s_1, s_2 \in \mathcal{S}$.
3. (Mereological order) $s_1 \sqsubseteq s_2 \iff \exists s_3 \in \mathcal{S}, s_1 \oplus s_3 = s_2$.
4. (Common part) there is a minimum $s_1 \sqcap s_2$ for $s_1, s_2 \in \mathcal{S}$.

We often call truthmakers *states*, because truthmakers represent “the state of affairs” in the world.

There is a natural model of axioms in Definition 2 as sets of *evidence*. Let \mathcal{E} be a set (of *evidence*).

Proposition 1 (Semantics of evidence). *Let \mathcal{S} be the power set of \mathcal{E} . Let 0 be the empty set \emptyset , \oplus be the set-theoretic union \cup , \sqsubseteq be the inclusion \subseteq , and \sqcap be the intersection \cap . Then, \mathcal{S} satisfies the axioms in Definition 2.*

Proof. (Axiom 1.) $\emptyset \cup s = s \cup \emptyset = s$. (Axiom 2.) $s_1 \cup s_2 = s_2 \cup s_1$. (Axiom 3.) Let $s_2 \setminus s_1$ be the set subtraction $s_2 \cap s_1^c$. Then, $s_1 \cup (s_2 \setminus s_1) = s_2$. (Axiom 4.) $s_1 \cap s_2$ is the minimum of s_1 and s_2 with respect to the order \subseteq .

In particular, the axioms of Definition 2 are consistent, because they have a model.

4.2 Frame, interpretation, and confidence

A *frame* determines how a GSN diagram is interpreted as an inference on truthmakers (in the state space \mathcal{S}). First, we define when two states are *orthogonal*.

Definition 3. *Two states s_1 and s_2 are orthogonal if $s_1 \not\sqsubseteq s_2$ and $s_2 \not\sqsubseteq s_1$. If all the elements of \mathcal{S} are mutually orthogonal, \mathcal{S} is said to be orthogonal as well.*

Let the set of evidence $\mathcal{E} \subseteq \mathcal{S}$ be orthogonal. Note that \mathcal{E} is different from \mathcal{E} in Proposition 1

Definition 4. *A tuple $\langle \mathcal{S}, \mathcal{E}, \mathbf{st} \rangle$ is called a frame.*

- \mathcal{S} is a state space.
- $\mathcal{E} \subseteq \mathcal{S}$ is the set of evidence.
- \mathbf{st} is a set of strategies, which are monotone functions from \mathcal{S} to \mathcal{S} .

We assume the following properties.

1. For any evidence e , strategy \mathbf{st} , and state s , $e \sqsubseteq \mathbf{st}(s)$ only when $e \sqsubseteq s$.
2. \mathbf{st} contains a special strategy \mathbf{id} called the identity strategy. \mathbf{id} is the identity function on the state space \mathcal{S} .

Independence of evidence means that all the evidence must be independently verified. We use this property to show that having multiple evidence increases confidence. Property 1 states that no strategy can infer the evidence unless that evidence is already verified.

Using a frame, we interpret a GSN diagram and its validity.

Definition 5 (Interpretation of a GSN diagram). We assume that a goal g is an element of \mathcal{S} such that $g \neq 0$ and evidence e is an element of \mathcal{E} .

$$\llbracket \langle g, \diamond \rangle \rrbracket \rho := 0 \quad (5)$$

$$\llbracket \langle g, e \rangle \rrbracket \rho := e \in \mathcal{E} \quad (6)$$

$$\llbracket \langle g, \text{OR}, (D_1, \dots, D_n) \rangle \rrbracket := \llbracket D_1 \rrbracket \sqcap \dots \sqcap \llbracket D_n \rrbracket \quad (7)$$

$$\llbracket \langle g, \text{st}, (D_1, \dots, D_n) \rangle \rrbracket := \text{st}(\llbracket D_1 \rrbracket \oplus \dots \oplus \llbracket D_n \rrbracket) \quad (8)$$

Definition 6 (Validity of a GSN diagram). Let D be a diagram. For a state $s \in \mathcal{S}$, we say that D justifies s whenever $s \sqsubseteq \llbracket D \rrbracket$. D is valid if the goal g of D is justified by D .

The distinction between inductive and deductive inferences [9] can be defined as follows.

Definition 7 (Inductive and deductive strategies). If a strategy $\text{st} \in \mathbf{st}$ satisfies $\text{st}(s) \supseteq s$ for all $s \in \mathcal{S}$, st is said to be inductive. If $\text{st} \in \mathbf{st}$ satisfies $\text{st}(s) \sqsubseteq s$ for all $s \in \mathcal{S}$, st is said to be deductive.

Definition 8. Confidence \mathcal{C} is any partial order. Confidence evaluation θ is a mapping from \mathcal{S} to \mathcal{C} such that for any $s_1, s_2 \in \mathcal{S}$, if $s_1 \sqsubseteq s_2$, then $\theta(s_2) \leq \theta(s_1)$. If, for any $s_1 \sqsubset s_2$, $\theta(s_2) < \theta(s_1)$ holds, then θ is said to be strict.

If $s_1 \sqsubseteq s_2$, s_2 has less confidence because it states more details about the state of the world compared to s_1 .

Theorem 1. If D is a valid GSN diagram and g is a goal of D , then $\theta(g) \geq \theta(\llbracket D \rrbracket)$.

Proof. Since $g \sqsubseteq \llbracket D \rrbracket$.

5 Relation to probabilistic evaluation

Probability is widely used for the confidence evaluation of GSN diagrams [4, 3, 6, 14, 13]. Although different methods and assumptions have been used to assign probability in the literature, they all satisfy the axioms of probability.

In this section, we show that our axioms are satisfied with any probabilistic evaluation with natural assumptions. Therefore, our axioms can be used to analyze the properties that hold for any probabilistic evaluation.

Theorem 2. Let $\langle \Omega, \mathcal{F}, P \rangle$ be a probability space, where Ω is the set of all samples, \mathcal{F} the set of all possible samples and P a probability measure on them. Then, \mathcal{F} can be regarded as a state space by $X \sqsubseteq Y \iff X \supseteq Y$, $X \oplus Y := X \cap Y$, $0 := \Omega$.

Proof. We check only axiom 3 in Definition 2. If $X \sqsubseteq Y$, then $X \supseteq Y$. Then, $X \oplus Y = X \cap Y = Y$. Conversely, if $X \cap Y = Y$, then $Y \sqsubseteq X$. Therefore, $X \sqsubseteq Y$.

Theorem 3. *A probability measure $P : \mathcal{F} \rightarrow [0, 1]$ is a confidence evaluation. If, for non-empty X , $P(X) > 0$ holds, then P is a strict confidence evaluation.*

Proof. If $X \sqsubseteq Y$, then $X \supseteq Y$. Therefore, $P(X) \geq P(Y)$. Further, if $X \sqsubset Y$, then $X \supset Y$. Therefore, there is a non-empty set Z such that $Y \cup Z = X$ and $Y \cap Z = \emptyset$. By the axiom of probability, $P(Y) + P(Z) = P(X)$. If $P(Z) > 0$, then $P(Y) < P(X)$.

We say that P is strict if, for any non-empty X , $P(X) > 0$.

Theorem 4. *Let E_1, \dots, E_n be independent (in the sense of probability theory) events that are not equal to Ω . If P is strict, then $\mathcal{E} = \{E_1, \dots, E_n\}$ forms a set of evidence.*

Proof. Let $A \setminus B = A \cap B^c$ be a set subtraction. First, note that $P(E_i) \neq 1$ because $\Omega \setminus E_i$ is non-empty. Assume that $E_1 \sqsubseteq E_2$. Then, $E_1 \supseteq E_2$ holds. Therefore, $P(E_1 \cap E_2) = P(E_2) \neq P(E_1)P(E_2)$ because $P(E_2) \neq 1$. This contradicts the independence of E_1 and E_2 .

Theorem 5. *Let \mathbf{st} be a set of monotone functions over 2^Ω such that $\mathbf{st} \in \mathbf{st}$ satisfies $X \supseteq \mathbf{st}(X)$ and $P(\mathbf{st}(X) \mid E) \leq P(X \mid E)$ for any evidence E . Then, $\langle \Omega, \mathcal{E}, \mathbf{st} \rangle$ is a frame.*

Proof. We only need to prove that $E \sqsubseteq \mathbf{st}(X)$ only when $E \subseteq X$. Assume that $E \sqsubseteq \mathbf{st}(X)$. Then, $P(\mathbf{st}(X) \mid E) = 1$ because $E \supseteq \mathbf{st}(X)$. Because $P(\mathbf{st}(X) \mid E) \leq P(X \mid E) = 1$, $E \subseteq X$.

6 Multi-legged argument

Bloomfield et al. [3] argued that multi-legged arguments can increase confidence. In this section, we present a sufficient condition for multi-legged arguments to increase confidence.

A multi-legged argument can be written using the OR construct, as shown in Fig. 2.

Theorem 6. *If $\llbracket D_1 \rrbracket$ and $\llbracket D_2 \rrbracket$ are independent and a confidence evaluation θ is strict, having a multi-legged argument increases confidence.*

Proof.

$$\theta(\llbracket D \rrbracket) = \theta(\llbracket D_1 \rrbracket \sqcap \llbracket D_2 \rrbracket) > \theta(\llbracket D_1 \rrbracket), \theta(\llbracket D_2 \rrbracket) \quad (9)$$

because $\llbracket D_1 \rrbracket \sqcap \llbracket D_2 \rrbracket \sqsubset \llbracket D_1 \rrbracket, \llbracket D_2 \rrbracket$.

The next theorem gives sufficient conditions of D_1 and D_2 , which makes Theorem 6 hold.

Theorem 7. *Suppose that the following conditions hold:*

- D_1 and D_2 contain only inductive inferences.

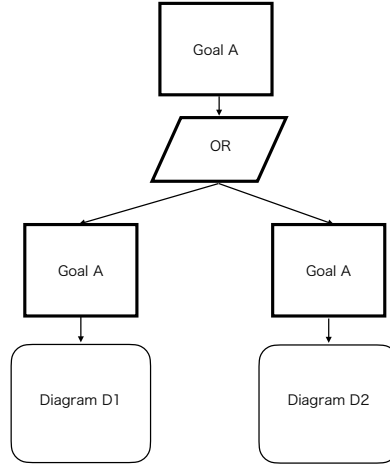


Fig. 2. Multi-legged argument D

- For set $E(D_1)$ of the evidence of D_1 and $E(D_2)$ of the evidence of D_2 , neither $E(D_1) \subseteq E(D_2)$ nor $E(D_2) \subseteq E(D_1)$ holds.
- D_1 and D_2 do not use a multi-legged argument.

Then, the multi-legged argument using D_1 and D_2 increases confidence.

Proof. There is evidence that $e \notin E(D_1)$ whereas $e \in E(D_2)$. By induction on D_1 , $e_1 \sqsubseteq \llbracket D_1 \rrbracket$. Here, we use the fact that D_1 has no multi-legged argument. If D_1 has a multi-legged argument, $e \sqsubseteq \llbracket D_1 \rrbracket$ may not hold. By property 1 of Def 4, we can show that $e \not\sqsubseteq \llbracket D_2 \rrbracket$ by induction on D_2 . Therefore, $\llbracket D_1 \rrbracket \not\sqsubseteq \llbracket D_2 \rrbracket$. By a similar argument, $\llbracket D_2 \rrbracket \not\sqsubseteq \llbracket D_1 \rrbracket$. Therefore, $\llbracket D_1 \rrbracket$ and $\llbracket D_2 \rrbracket$ are independent. By Theorem 6, we obtain the conclusion of the theorem.

7 Concrete example

In this section, we analyze a part of the concrete GSN diagram shown in Fig. 3, which argues the correctness of a computer simulation of a biological process, namely lymphoid tissue formation. Using our framework, we can clarify the nature of arguments and suggest further improvement.

Fig. 3 shows the argument for claim 1.1.4: “simulation captures cell aggregation emergent behavior at 72 h.” Claim 1.1.4 is derived from three strategies; thus, we can regard the argument as a multi-legged argument. As discussed in Section 6, if each argument is based on different sets of evidence, contains only inductive strategies, and does not contain another multi-legged argument, then a multi-legged argument improves confidence.

¹ Licensed under the Creative Commons Attribution 4.0 International License

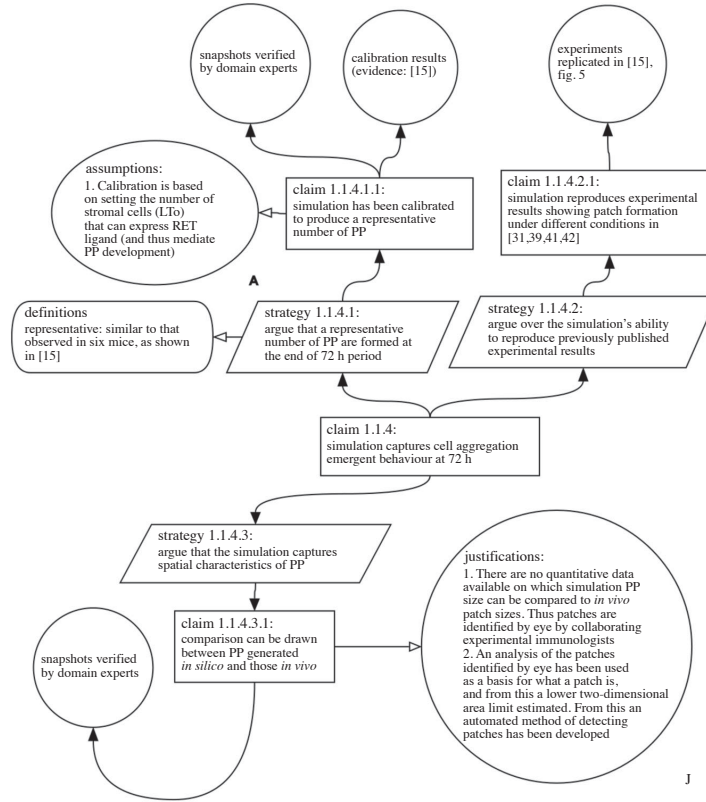


Fig. 3. GSN used for justification of a biological simulation (claim 1.1.4) [1] ¹

However, claim 1.1.4.3.1 in Fig. 3 uses the opinions of domain experts as the only evidence. Because claim 1.1.4.1.1 also uses the opinions of domain experts, the evidence used in claim 1.1.4.3.1 might be contained in that of claim 1.1.4.1.1, which violates the condition presented in Section 6. Therefore, the multi-legged argument based on claims 1.1.4.1.1 and 1.1.4.3.1 may not increase the confidence of claim 1.1.4. An explicit description of the opinion of domain experts would increase the confidence of the argument by differentiating the evidence used by claims 1.1.4.1.1 and 1.1.4.3.1.

8 Conclusion and Future work

This paper presented a framework for interpreting and evaluating assurance cases in an abstract manner. Unlike numerical evaluation methods, our method is purely qualitative, which we consider an advantage because the reasoning of assurance cases is inherently qualitative. We demonstrated that our method can provide a rigorous theoretical foundation for the qualitative evaluation of assur-

ance cases, using multi-legged arguments (Section 6) and a concrete case study (Section 7) as examples. Furthermore, we showed that probabilistic evaluations are special cases of our method in Section 5.

In the future, we plan to investigate additional axioms to realize a finer confidence evaluation. Further, we plan to investigate the relation of our method to other methods, especially the Dempster–Shafer theory.

References

1. Alden, K., Andrews, P.S., Polack, F.A., Veiga-Fernandes, H., Coles, M.C., Timmis, J.: Using argument notation to engineer biological simulations with increased confidence. *Journal of the Royal Society Interface* **12**(104), 20141059 (2015)
2. Assurance Case Working Group: Goal structuring notation community standard version 2 (January 2018), downloaded from <https://scsc.uk/r141B:1>
3. Bloomfield, R., Littlewood, B.: Multi-legged arguments: The impact of diversity upon confidence in dependability arguments. *Proceedings of the International Conference on Dependable Systems and Networks* (June 2014), 25–34 (2003)
4. Bloomfield, R.E., Littlewood, B., Wright, D.: Confidence: Its role in dependability cases for risk assessment. In: *The 37th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, DSN 2007, 25-28 June 2007, Edinburgh, UK, Proceedings*. pp. 338–346. IEEE Computer Society (2007)
5. Fine, K.: Truthmaker Semantics. *A Companion to the Philosophy of Language* (February 2017), 556–577 (2017)
6. Guiochet, J., Hoang, Q.A.D., Kaâniche, M.: A model for safety case confidence assessment. In: Koornneef, F., van Gulijk, C. (eds.) *SAFECOMP 2015 Delft, The Netherlands, September 23-25, 2015. Proceedings. Lecture Notes in Computer Science*, vol. 9337, pp. 313–327. Springer (2015)
7. Maksimov, M., Kokaly, S., Chechik, M.: A survey of tool-supported assurance case assessment techniques. *ACM Computing Surveys* **52**(5) (2019)
8. Matsuno, Y.: A Design and Implementation of an Assurance Case Language. *Proceedings - 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, DSN 2014* pp. 630–641 (2014)
9. Rushby, J.: Assurance and assurance cases. In: Pretschner, A., Peled, D., Hutzelmann, T. (eds.) *Dependable Software Systems Engineering (Marktoberdorf Summer School Lectures, 2016)*, pp. 207–236. Volume 50 of NATO Science for Peace and Security Series D, IOS Press (Oct 2017)
10. The MISRA consortium: Guidelines for automotive safety arguments (2019)
11. Varzi, A.: Mereology. In: Zalta, E.N. (ed.) *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2019 edn. (2019)
12. Wang, R., Guiochet, J., Motet, G., Schön, W.: Modelling confidence in railway safety case. *Safety Science* **110**(December), 286–299 (2018)
13. Wang, R., Guiochet, J., Motet, G., Schön, W.: Safety case confidence propagation based on dempster-shafer theory. *Int. J. Approx. Reason.* **107**, 46–64 (2019)
14. Weinstock, C.B., Goodenough, J.B., Klein, A.Z.: Measuring assurance case confidence using baconian probabilities. In: *Proceedings of ASSURE '13, San Francisco, 2013*. pp. 7–11. IEEE Computer Society (2013)