Quelle argumentation pour des systèmes de confiance ?

Jean-Michel Bruel, IRIT, Université de Toulouse, Toulouse
Rémi Delmas, Uber Advanced Technology Center, Paris
Régine Laleau, Université Paris-Est Créteil, LACL, Créteil
Thomas Polacsek, ONERA, Toulouse
Florence Sedes, IRIT, Université de Toulouse, Toulouse

1 Contexte

La nécessité d'avoir confiance dans le fonctionnement d'un logiciel est quelque chose de connu, étudié et pratiqué depuis de nombreuses années dans, par exemple, le monde de l'aéronautique, du ferroviaire ou du médical. Dans le cadre d'applications critiques, au sens où des vies humaines sont en jeu, il est nécessaire de s'assurer qu'un logiciel fonctionne correctement, c'est-à-dire conformément à ce que l'on attend de lui. Cette exigence de confiance n'est cependant pas circonscrite aux seuls domaines impactant des vies humaines, nous la retrouvons aussi dans des domaines sensibles, comme le domaine bancaire ou la distribution d'énergie. Nous parlerons donc ici d'informatique de confiance. Notons que la confiance n'est pas un concept absolu, on a confiance en quelque chose pour quelque chose. Dès lors, si nous considérons le besoin de confiance dans une application pour quelque chose de précis, nous pouvons étendre le domaine de l'informatique de confiance à toute application qui doit absolument garantir un certain fonctionnement. En plus des applications critiques, nous pouvons inclure les applications qui gèrent des données privées dans lesquelles nous voulons avoir confiance en ce qui concerne la protection des données ou les applications de type blockchain et contrats intelligents dans lesquelles nous voulons avoir confiance en ce qui concerne l'impossibilité de manipulations frauduleuses.

Dans une optique d'établir un haut niveau de confiance, nous avons vu, depuis de nombreuses années, le développement et l'usage des méthodes formelles pour garantir des propriétés sur des systèmes. Cependant, l'utilisation de méthodes formelles ne nous ôte pas de certains doutes. Considérons que nous disposons de la preuve mathématique de la correction d'un artefact, sommes-nous sûrs que cette preuve ne contienne pas elle-même des erreurs ? Si elle a été établie par une machine, sommes-nous sûrs que le programme utilisé est lui aussi prouvé ? Nous pouvons ainsi remettre en question tous les éléments, chercher des preuves aux preuves, sans jamais trouver de fin à nos questionnements. Nous sommes typiquement face au problème épistémologique de la régression infinie. Loin d'être un problème purement philosophique, le problème de la confiance dans les moyens utilisés pour établir la preuve de correction se pose cruellement dans le monde de l'ingénierie en général et, plus particulièrement, dans le cadre de la certification. En effet, toute la démarche visant à certifier un artefact n'a qu'un seul but : prévenir les erreurs. Il est donc crucial que les moyens utilisés ne soient pas eux-mêmes entachés d'erreurs ou, tout au moins, que nous ayons confiance en eux.

Nous pouvons simplifier la certification en considérant qu'il s'agit de « savoir si un artefact est correct ». Dans ce contexte, la preuve formelle n'est qu'un élément parmi d'autres permettant d'établir cette connaissance. Comme le souligne Tony Hoare [4], ce n'est pas grâce à l'utilisation de méthodes formelles que les logiciels sont devenus plus fiables, mais par l'usage de techniques déjà employées

dans les autres branches de l'ingénierie comme : des procédures rigoureuses de relecture des spécifications de conceptions, l'assurance qualité fondée sur de larges éventails de tests ou de l'amélioration continue. Dès lors, comment, à partir de ces éléments informels, être sûr que l'artefact final est correct ?

Le problème qui nous préoccupe ici est en fait un problème d'inférence. Nous cherchons à déterminer s'il est acceptable ou pas de passer d'un ensemble de justifications à une conclusion. Pour être plus précis, nous visons l'étude des documentations techniques qui permettent la certification d'artefacts. Nous trouvons ce type de documents, par exemple, dans le domaine de l'évaluation des risques et de la fiabilité des systèmes critiques sous le nom de safety case ou d'assurance case. Le safety case est un document structuré qui fournit une justification et des arguments valables sur le fait qu'un système satisfait des propriétés relatives à sa sécurité.

Parallèlement, hors de la sphère de la certification, depuis quelques années, nous voyons émerger le besoin de démontrer la conformité d'un système par rapport à une norme. En effet, qu'il s'agisse de sécurité ou de protection des données et de la vie privée, les systèmes doivent de plus en plus se conformer à un nombre croissant de réglementations. Ce besoin de conformité à de nouvelles règles (comme par exemple le nouveau règlement général sur la protection des données de l'Union Européenne) implique des changements techniques profonds. Ces règles doivent non seulement être prises en compte par les systèmes, mais il est également nécessaire de démontrer qu'un système s'y conforme. Outre la protection de la vie privée, nous pouvons légitimement penser que, dans le futur, les systèmes, qu'ils soient déjà existants ou à construire, devront de plus en plus démontrer leur conformité à des attentes plus sociétales telles que le développement durable, l'éco-énergétique (energy-aware programming) ou des considérations d'explicabilité ou de traitements éthiques.

Par conséquent, nous devons prendre en compte deux aspects : premièrement, veiller à ce qu'un système soit conforme à un règlement et, deuxièmement, faire en sorte que les moyens de mise en conformité soient accessibles à tous.

Si nous nous basons sur les pratiques existantes dans l'univers de la certification, une grande partie de la démonstration de conformité, ce qui permet d'établir la confiance, est basée sur une approche processus et sur les aspects organisationnels. Ainsi, une organisation doit démontrer que sa structure, ses rôles et processus sont conformes à ce qui est demandé : le fonctionnement effectif de l'organisation est conforme à la structure déclarée, les rôles sont pourvus et les différents acteurs sont conscients des missions exigées par les rôles.

Concernant les artefacts techniques, la réponse est bien évidemment moins organisationnelle, mais repose sur une approche de confiance par conception, design, au sens où la composante critique est prise en compte à l'origine. Cette approche de garantie par construction ne tient malheureusement plus une fois sorti du domaine critique. Pour des problèmes de coûts, ou tout simplement par ce que l'on cherche à montrer a posteriori, il serait illusoire de penser pouvoir appliquer une approche garantie par construction pour la conformité à des normes futures, surtout si celles-ci concernent des aspects non cruciaux pour le fonctionnement du système (comme l'éco-énergétique). De plus, cette approche extrêmement coûteuse de garantie par construction semble atteindre ses limites avec, par exemple, les problèmes soulevés par l'*embarquabilité* de l'apprentissage automatique.

2 Problématique

Ce défi se positionne à la convergence de ces deux problématiques que sont la certification et la conformité à un règlement ou des attentes sociétales. Dans les deux cas, il est nécessaire d'identifier les exigences propres à ce besoin de convaincre une autorité ou des usagers. En ce qui concerne le res-

pect d'une réglementation, pour faire valoir qu'un système s'y conforme il faut tout d'abord identifier les exigences résultant de cette réglementation. De plus, dans le contexte d'un système préexistant, il peut être utile d'effectuer une étude d'impact réglementaire sur un système. Il est à noter que cette articulation entre le texte d'un règlement (une norme, une loi, etc.), qui tend à définir des objectifs de haut niveau, et un système est commun dans le monde de la certification, qui est également basé sur des normes. Par ailleurs, il est aussi important de s'intéresser à la structuration de l'argumentation qu'elle relève de la certification ou de la conformité à un règlement. Dans les deux cas, en portant notre attention sur les aspects justification, nous opérons un glissement du raisonnement déductif, de la preuve logique, vers une forme de raisonnement plus informel. Cet aspect informel ne doit pas nous arrêter dans notre démarche. En effet, il doit être possible de dégager des structures permettant de capturer la rationalité de telles argumentations et de définir des approches pour prévenir les raisonnements fallacieux.

3 Challenges identifiés

Parmi les travaux qui cherchent à organiser une argumentation dans le but de convaincre de l'efficience d'un système, nous pouvons citer tout ce qui se rapproche de près ou de loin de la thématique des *assurance cases*. Un assurance case est "une argumentation structurée selon laquelle un système est acceptable pour l'usage auquel il est destiné en ce qui concerne des préoccupations spécifiques" [1] [10]. Si, en pratique, il n'y a pas d'exigence particulière sur le format et la structuration d'un assurance case, de nombreux travaux proposent de structurer l'ensemble des justifications sous une forme inspirée du schéma de Toulmin [12]. Parmi ces approches, nous pouvons citer par exemple *Goal Structuring Notation* (GSN) [5, 6], *Claim-Argument-Evidence* [2], une approche textuelle de John Rushby [11] *Justification Diagram* [9] ou *Structured Assurance Case Meta-model* [8]. Dans cette idée de structuration, tout reste à faire. D'ailleurs des travaux récents suggèrent qu'une approche par patrons de conception pourrait-être une solution possible [1, 3, 7, 13].

Par ailleurs, puisqu'il est question d'argumentation et de justification, des liens peuvent être créés avec des groupes de travail du GDR GPL (MFDL, IE, GLACE, MTV2, IDM, AFSEC, ...) et des GDR qui s'intéressent à la question : Sécurité Informatique, Réseaux et Systèmes Distribués (RSD), Masses de Données, Informations et Connaissances en Sciences (MaDICS), Aspects Formels et Algorithmiques de l'Intelligence Artificielle (IA), Traitement Automatique des Langues (TAL)

Comme nous le voyons, beaucoup de travaux restent à mener et bien des questions méritent d'être creusées. Notre défi s'intéressera en particulier aux challenges suivants :

- Comment exprimer les exigences relatives à une norme, un standard, une réglementation ?
- Comment relier ces exigences aux exigences fonctionnelles et non fonctionnelles (comme la sécurité et la sûreté) d'un système ?
- Comment s'assurer qu'un système est conforme à une réglementation ?
- Comment un système d'information peut aider un processus d'argumentation ?
- Comment renforcer, d'un point de vue argumentatif, la complémentarité entre preuves, tests et simulations ?
- Quels formalismes d'argumentation et de méthodes pour une approche incrémentale de la mise au point et de la certification des systèmes ?
- Modularité, réutilisation, analyses d'impact ?

1. Traduction de "*an organized argument that a system is acceptable for its intended use with respect to specified concerns*".

— La plupart des outils existants d'argumentation sont graphiques et difficilement utilisables pour des gros projets industriels, comment développer des interfaces utilisateurs pour la saisie et la navigation d'argumentations complexes ?

Références

[1] Clément Duffau, Thomas Polacsek, and Mireille Blay-Fornarino. Support of justification elicitation : Two industrial reports. In *Proceedings of International Conference Advanced Information Systems Engineering, CAiSE 2018*, 2018.

[2] Luke Emmet and George Cleland. Graphical notations, narratives and persuasion : a pliant systems approach to hypertext tool design. In *Proceedings of Hypertext and Hypermedia, HYPERTEXT 2002*, 2002.

[3] Richard Hawkins, Tim Kelly, John Knight, and Patrick Graydon. A new approach to creating clear safety arguments. In *Advances in systems safety*. Springer, 2011.

[4] C. Hoare. How did software get so reliable without proof ?, 1996. `https://www.gwern.net/docs/math/1996-hoare.pdf`. Last Accessed : 1-4-2020.

[5] Tim Kelly and Rob Weaver. The goal structuring notation - a safety argument notation. In *DNS 2004 Workshop on Assurance Cases*, 2004.

[6] John A McDermid. Support for safety cases and safety arguments using SAM. *Reliability Engineering & System Safety*, 43(2), 1994.

[7] Dominique Méry, Bernhard Schätz, and Alan Wassyng. The pacemaker challenge : Developing certifiable medical devices (Dagstuhl seminar 14062). In *Dagstuhl Reports*, volume 4 :2, 2014.

[8] OMG. Structured assurance case meta-model (SACM). Technical report, Object Management Group, 2013.

[9] Thomas Polacsek. Validation, Accreditation or Certification : a New Kind of Diagram to Provide Confidence. In *Proceedings of International Conference on Research Challenges in Information Science, RCIS*, 2016.

[10] David J Rinehart, John C Knight, and Jonathan Rowanhill. Current practices in constructing and evaluating assurance cases with applications to aviation. Technical report, NASA, 2015.

[11] John Rushby, Xidong Xu, Murali Rangarajan, and Thomas L Weaver. Understanding and evaluating assurance cases. Technical Report NASA/CR-2015-218802, NASA Langley Research Center, 2015.

[12] Stephen E. Toulmin. *The Uses of Argument*. Cambridge University Press, Cambridge, UK, 2003. Updated Edition, first edition 1958.

[13] Alan Wassyng, Paul Joannou, Mark Lawford, Maibaum Thomas, and Neeraj Kumar Singh. New standards for trustworthy cyber-physical systems. In *Trustworthy Cyber-Physical Systems Engineering*, chapter 13, pages 337–368. Addison-Wesley Longman Publishing, 2016.