

HORIZON
2020

Devising certifiable and explainable algorithms for verification and planning in cyber-physical systems

Results in Brief

Explaining an algorithm's output

Before algorithms can be used in critical operations like air traffic control, human users must be able to trust them. Explanations, called certificates, on why an algorithm's output is correct, could serve as a foundation for building this trust.



DIGITAL ECONOMY



© Asset data, Shutterstock

Algorithms are a set of rules or instructions that tell a computer what to do. While these processes play an essential role helping software developers create efficient and error-free programmes, they do have their limitations – one of the main ones being they often cannot provide any tangible proof of their results.

According to [Shaul Almagor](#), an assistant professor of Computer Science at [Technion](#), this lack of certification presents a real challenge, especially in designing the cyber-physical systems of tomorrow.

“Take for example a robotic planning algorithm, where a robot is asked to undertake a task like finding a key and opening a door,” he explains. “Now suppose that the algorithm tells the user that no such plan exists for the robot – how can the user know this answer is correct?”

Or what if the algorithm does find a plan, but it is very complicated. How can the user

trust the plan? “This problem is extremely important in settings where a human supervisor must approve a safety-critical operation, such as in air traffic control or in an industrial factory,” says Almagor. “In such settings, it would be ideal if the algorithm could provide a simple explanation on why its output is correct.”

With the support of the EU-funded ALGOCERT project, Almagor set out to devise such explanations, which he calls certificates.

Tailoring certificates to specific contexts

As Almagor explains, to be meaningful, these certificates must be tailored to specific contexts. “In order to define meaningful notions of certificates, we must build upon the specific properties of the planning task,” he adds. “Only then can we attempt to devise certifiable algorithms.”

Take for instance the reachability problem for linear [dynamical systems](#) , which represent the evolution in time of vectors. Although these systems use simple mathematical rules, they can give rise to very complex behaviours.

“If the evolution reaches a target, this is easy to certify – just provide the reaching time,” explains Almagor. “But if it doesn’t reach the target, it’s not clear how to give a certificate.”

Based on this, the project proposed a notion of certificates for non-reaching scenarios. “We synthesise a set that contains the entire evolution of the system but that does not intersect the target,” adds Almagor.

Another problem addressed by the project is that of multi-agent pathfinding, or MAPF. According to Almagor, MAPF represents a fundamental problem in robotic planning, including such robotic applications as automated warehouses and vehicles.

“These applications require multiple agents being able to concurrently follow established paths without colliding into each other,” notes Almagor.

Here the project proposed a scheme that explains the correctness of an MAPF plan to a human user. “We convince the user that the plan has no risk of collision by decomposing it into a sequence of images, with the paths of the agents being disjointed in each image,” Almagor says. “This really is an intuitive way of explaining non-collision.”

An important stepping stone

These are just a few of the problems addressed during the [Marie Skłodowska-Curie Actions](#) supported ALGOCERT project, many of which have been fully explained in various papers and presentations.

“The notion of explainability has gained a lot of attention over the past few years,” concludes Almagor. “Although I can’t say that our methods are the holy grail of algorithm certification, I do believe they represent an important stepping stone along the way.”

Keywords

[ALGOCERT](#)

[algorithms](#)

[explanations](#)

[certificates](#)

[certification](#)

[software](#)

[cyber-physical systems](#)

[robot](#)

[safety-critical operation](#)

[dynamical systems](#)

[mathematical](#)

[multi-agent pathfinding](#)

Project Information

ALGOCERT

Grant agreement ID: 837327

[Project website](#)

DOI

[10.3030/837327](#)

Project closed

EC signature date

29 April 2019

Start date

15 July 2019

End date

14 July 2021

Funded under

EXCELLENT SCIENCE - Marie Skłodowska-Curie Actions

Total cost

€ 185 464,32

EU contribution

€ 185 464,32

Coordinated by

TECHNION - ISRAEL INSTITUTE
OF TECHNOLOGY

 Israel

Last update: 10 December 2021

Permalink: <https://cordis.europa.eu/article/id/435398-explaining-an-algorithm-s-output>

European Union, 2025

