Towards a Code of Ethics for Autonomous and Self-Adaptive Systems

Danny Weyns
danny.weyns@kuleuven.be
KU Leuven Belgium & Linnaeus University Sweden

ABSTRACT

Software systems are playing an increasingly important role in many domains of our society. To ensure that software will support the public good, software engineers, who create and maintain the software, shall adhere to ethical principles. A joint task force of IEEE and ACM has brought such a set of principles together in a Code of Ethics. These principles describe responsibilities for software engineers and guidelines to assist them when making decisions in the benefit of public good. With the emergence of computing systems that take autonomous decisions, there is growing consensus that new ethical principles will be required. Since self-adaptive systems are characterized by autonomy, the need for new principles applies to these systems. Based on the Code of Ethics and leveraging on ongoing initiatives, we suggest an initial set of new ethical principles for autonomous and self-adaptive systems as an inspiration for an extended Code of Ethics for this important class of systems.

KEYWORDS

Autonomy, self-adaptation, ethical principles, code of ethics

ACM Reference Format:
Danny Weyns. 2020. Towards a Code of Ethics for Autonomous and Self-Adaptive Systems. In *SEAMS '20: 15th International Symposium on Software Engineering for Adaptive and Self-Managing Systems, May 25–26, 2020, Seoul, South Korea.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/1122445.1122456

1 MOTIVATION

Software systems are playing an increasingly important role in industry, government, health, transportation, education, and many other domains of our society. Because software is created and maintained by software engineers, they have significant opportunities to do good or cause harm, directly or indirectly. To ensure that software will be used for good and make the world a better place, software engineers shall adhere to ethical principles.

A joint task force of IEEE and ACM has brought these principles together in a "Code of Ethics" [5, 8]. These principles describe responsibilities and guidelines for software engineers and define what is good and right behavior when making decisions. With the emergence of computing systems that take decisions autonomously, usually on behalf of stakeholders, new additional ethical principles will be required, as endorsed by several initiatives all over the world. Since self-adaptive systems are characterized by autonomy, i.e., the system adapts itself with some level of autonomous behavior [1, 6, 9], the need for new ethical principles applies to these systems.

Based on the Code of Ethics and leveraging on ongoing initiatives, we suggests a set of new ethical principles for autonomous and self-adaptive system, and we reflect on how to incorporate these new principles in an extended Code of Ethics for these systems.

This paper is structured as follows. Section 2 provides a brief introduction to the Code of Ethics. In Section 3, we suggest a number of new ethical principles for autonomous systems. Finally, Section 4 reflects on implementing the new principles to establish an extended Code of Ethics for autonomous and self-adaptive systems.

2 CODE OF ETHICS

IEEE and ACM's Code of Ethics describes how basic ethical principles apply to software professionals and serves as a basis for ethical decision-making. Figure 1 gives an overview of Code of Ethics.

The items numbered from 1 to 8 refer to the ethical principles that are defined in the Code of Ethics. We briefly summarize these

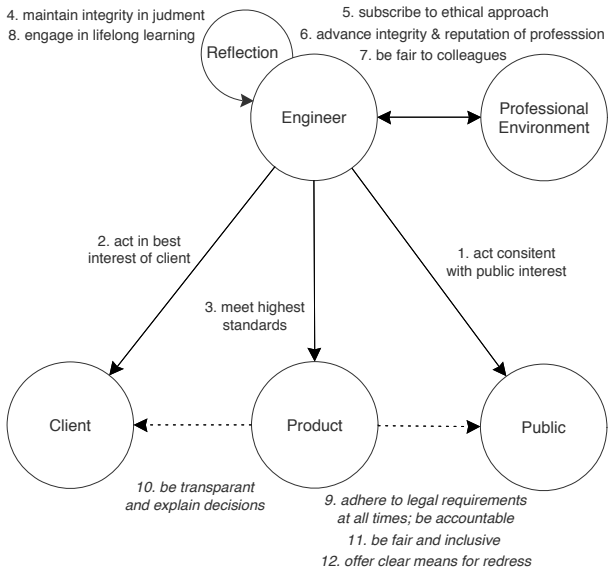

Figure 1: Schematic overview of the Code of Ethics of IEEE/ACM; circles represent actors, arrows represent ethical principles, and dotted arrows suggest new ethical principles for autonomous systems, such as self-adaptive systems.

SEAMS '20, May 25–26, 2020, Seoul, South Korea

ACM ISBN AAA-B-CCC-XXXX-X/YY/ZZ…$XX.YY
https://doi.org/10.1145/1122445.1122456

principles. Principle 1 states that software engineers shall act consistently with the public interest. Principle 2 states that software engineers shall act in a manner that is in the best interests of their clients, consistent with the public interest. The third principle states that software engineers shall ensure that their products and related modifications meet the highest professional standards possible. Principle 4 states that software engineers shall maintain integrity and independence in their professional judgment, while principle 5 states that software engineering professionals shall subscribe to and promote an ethical approach to the management of software development and maintenance. Principle 6 states that software engineers shall advance the integrity and reputation of the profession consistent with the public interest. Principle 7 states that software engineers shall be fair and supportive to their colleagues. Finally, principle 8 states that software engineers shall participate in lifelong learning regarding the practice of their profession. While some of these principles may need to be re-interpreted in terms of autonomy and self-adaptation, our focus here is on new ethical principles specifically for autonomous and self-adaptive systems.

3 ETHICAL PRINCIPLES FOR AUTONOMOUS AND SELF-ADAPTIVE SYSTEMS

For computing systems that take autonomous decisions, new additional ethical principles will be required. Several initiatives all over the world have been investigating the impact of the autonomy of software systems to make decisions on the ethical principles. For example, the United States' Fair Credit Reporting Act and European Union's General Data Protection Regulation (GDPR) prescribe that data must be processed in a way that is fair and unbiased. GDPR also alludes to the right of an individual to receive an explanation about decisions made by an automated system. Another example is the recommendation provided by The World Economic Forum to prevent discriminatory outcomes by machine learning applications [4]. Researchers also take initiatives towards the definition of ethical frameworks, for instance for the design of autonomous intelligent systems [7]. Since self-adaptive systems adapt autonomously (or with only the required human intervention) to deal with changing conditions, the need for new principles applies to these systems.

The items numbered from 9 to 12 in Figure 1 suggest an initial set of new ethical principles for autonomous systems. Whereas the basic principles (1 to 8) directly apply to the engineers of software systems, the new principles are exposed through the systems themselves, and apply indirectly to the engineers that developed the systems. Hence it is the responsibility of the engineer to ensure that the system complies to all new principles. For principles 9 and 10, an agreement grows among different actors. Principles 11 and 12 on the other hand are more speculative and subject of debate.

Legal Compliance. Principle 9 states that the system shall adhere to the legal requirements at all times and be accountable for that. Consider for instance an IoT-based ecosystem that supports food security for people in a city that buy goods in supermarkets [2]. Users may share information about grocery needs through a platform. Supermarkets can gather this data to better manage their stock and provide users with targeted promotions. Principle 9 requires that engineers that realize the system should ensure that the privacy of the data that users share should be respected at any time.

Transparency and Explainability. Principle 10 states that the system shall be transparent and explain the decisions it makes in a way that they are understandable to those that are affected by the decision-making. E.g., the food security system may collect data on behalf of the supermarkets and use that data to adapt the prices depending on the demand. Principle 10 requires that engineers build the system such that it allows users to understand what data is being collected and how this data is being used, enabling them to give informed consent for automatic data collection, get access to it, and delete if required.

Fairness and Inclusiveness. Principle 11 states that the system shall be fair and explicitly take fairness into account in its evaluation metrics. E.g., in the food security system, engineers of the system should ensure that the policies that are used to provide targeted promotions foster fair and inclusive participation of all users and do not discriminate on any ground.

Means for Redress. Principle 12 states that the system shall offer clear means for redress for those affected by disparate impacts of it. E.g., engineers of the food security system should ensure that the system offers users the means for redress when they suspect that their data is misused to manipulate the price of products.

4 ESTABLISHING A CODE OF ETHICS FOR SELF-ADAPTIVE SYSTEMS

We offer the initial set of ethical principles as inspiration for extending the Code of Ethics for autonomous and self-adaptive systems. Establishing such an extended code is complex and will require a huge concerted effort from a broad range of actors. Not only is there a need for an agreement on new ethical principles and possibly a re-interpretation of existing principles in terms of autonomy and self-adaptation, these principles need to be implemented, which raises numerous technical challenges. For instance, approving that autonomous and self-adaptive systems are safe and meet their specifications is particularly challenging since such systems operate under uncertainty that is often difficult to anticipate [3]. When a code ultimately expresses the consensus of a profession community on ethical issues, it may be supported by standards. This will then offer a means to educate the public that will be affected by autonomous software systems, but also teach new generations of software engineers about the ethical obligations in their profession.

REFERENCES

[1] J. Andersson et al. 2009. Modelling Dimensions of Self-adaptive Software Systems. In *Software Engineering for Self-Adaptive Systems*. LNCS, Vol. 5525. Springer.
[2] A. Bennaceur et al. 2016. Feed Me, Feed Me: An Exemplar for Engineering Adaptive Software. In *Software Engineering for Adaptive and Self-Managing Systems*. ACM.
[3] R. Calinescu et al. 2018. Engineering Trustworthy Self-Adaptive Software with Dynamic Assurance Cases. *Transactions on Software Engineering* 44, 11 (2018).
[4] World Economic Forum. 1/2020. https://www.weforum.org/whitepapers/how-to-prevent-discriminatory-outcomes-in-machine-learning/
[5] D. Gotterbarn et al. 2001. Software Engineering Code of Ethics and Professional Practice. *Science and Engineering Ethics* 7 (2001).
[6] J. Kephart and D. Chess. 2003. The Vision of Autonomic Computing. *Computer* 36, 1 (2003).
[7] J. Leikas et al. 2019. Ethical Framework for Designing Autonomous Intelligent Systems. *Journal of Open Innovation: Technology, Market, & Complexity* 5, 1 (2019).
[8] IEEE/ACM Joint Task Force on Software Engineering Ethics and Professional Practices. 1/2020. Code of Ethics. www.computer.org/education/code-of-ethics
[9] D. Weyns. 2019. Software Engineering of Self-adaptive Systems. In *Handbook of Software Engineering*. Springer, 399–443.