BRIEFING NOTES

BN-67-Emerging technology and military application-Aug2021

CHALLENGES OF PUBLIC POLICY IN AUTONOMOUS SYSTEMS

Authors: Shahram Shahkar[1] and Kash Khorasani[2]
1 Graduate student, Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada
2 Professor, Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada

SUMMARY

- As human command diminishes in autonomous systems there need to be some form of ethics to replace human conscience in the decision-making process of computer systems.
- The most important challenge in successful application of autonomy is explainability of decisions and how they have been made.
- Explainability of decisions is crucial for autonomy given that computers are tasked to undertake decisions that are highly complex and challenging to fully represent by human cognitive process.
- Since ethical considerations are challenging to formulate in autonomous systems is it possible to replace ethics and morality by social norms and values. If yes, how does that help public acceptance?

CONTEXT

- Autonomy is crucial in the future development of our civilization, to optimize and equally share our existing resources and infrastructure [2].
- Ethics in Autonomy is also essential given that human oversight is fully or partially removed or minimized. In compensation, an artificial form of ethics has to be substituted in order to maintain moral standards in our civilization [3].
- As human oversight gradually diminishes, machine decision making processes have to be understood and accepted among the users. Therefore, public opinion has to be taken into consideration through the development stages of these systems [1].
- It is challenging to formulate ethics in computers since computer operations are based on numeric quantities rather than qualitative parameters, and in order to follow ethical standards all alternative options in a given decision have to be quantized and then compared on order to deduce the most preferable choice.
- Social norms are more feasible to implement in computer systems since computers can learn repeated behaving patterns and replicate these patterns under given circumstances.
- Social norms are also quite explainable and favour public acceptance. Therefore, autonomous systems that follow social norms are expected to be acceptable and approved in societies.

FEASIBILITY

- One should investigate to determine if social norms are equivalent to the ethical standards. Only then one could substitute ethics by social norms in autonomous systems.

- Social norms represent behaviors that are culturally approved and/or widely accepted. Morals on another hand can be understood as distinguishing criteria for a behavior to be classified into as right or wrong regardless of whether or not the said behavior is culturally approved or widely accepted. Ethics is then the systematic approach one uses in making a decision regarding a behavior [4], [5].

- Values on another hand comprise of concepts and ideas that are preferred. In other words, what is good, right, wise or beneficial. According to [8] "Values are implanted early in a person's life and once they are fixed, serve as a guide in choosing behaviour and in forming attitudes". "Values account for the stability of social order, and they provide the general guidelines for social conduct".

- Consequently, values can be considered as standards of social behaviour that are derived from social interactions and are accepted as constituent facts of social structure. In this sense it could be possible to consider them equivalent to ethics in certain context [6],[7],[8].

- Since machines lack power of judgment, it appears that promotion of ethical machines would require quantitative comparison of available options for a decision to be made. A simpler approach could be expecting machines to learn and apt to social norms instead. This is due to the fact that it might be far more practical to devise a learning algorithm for a machine to learn from repeatedly observed behavioral patterns, referred to as social norms. The question however would be how closely social norms could replicate ethical behaviors, and if it is possible to substitute social norms with ethics in societies. In order to answer this question, let us first take a look at the most dominant definitions of social norms.

- Legal scholars have touted social norms as efficient alternatives to legal rules, as they may internalize negative externalities and provide signaling mechanisms at little or no cost (Ellickson 1991; Posner 2000).

- According to the Stanford's encyclopedia of philosophy [4], "social norms like many other social phenomena, are the unplanned result of individuals' interactions. It has been argued that social norms ought to be understood as a kind of grammar of social interactions. Like a grammar, a system of norms specifies what is acceptable and what is not in a society or group. And, analogously to a grammar, it is not the product of human design".

- Some authors ascribe norms with recurring patterns of behavior. Others relate them with normative beliefs and social expectations. According to Thomas Schelling [5] the norms that are most interesting to study are those that emerge naturally from individuals' interactions. Therefore, it might be prudent to analyze the conditions under which such norms come into existence. It is an agreeable fact among all the activists that social norms produce conformity and social order and that there exists a strong correlation between people's beliefs about the behaviors accepted and approved in the society and their behaviors. It is said that "Normative beliefs are habitually accompanied by the

expectation that other people will follow the prescribed behavior and avoid the proscribed one" [4].

+ The authors in [6],[7] have quoted that "the rational choice model of conformity maintains that, since norms are upheld by sanctions, compliance is a utility-maximizing strategy. Provided that conformity to a norm attracts approval and transgression disapproval, conforming is the rational thing to do, since nobody willfully attracts discredit and punishment". Therefore, "conforming behavior is rationally chosen in order to avoid negative sanctions or to attract positive sanctions" [4].

RELATIONSHIP BETWEEN NORMS AND VALUES

+ Norms and values have salient relationship. Norms are specific, whereas values are not. There may be, in a particular situation, delusion of norms, but values are commanding [8].

+ Norms are rules for behaving. They say more or less specifically what should or should not be done by particular types of actors in given circumstances.

+ Values are standard of desirability that are more nearly independent of specific situations. The same value may be a point of reference for a great many specific norms; a particular norm may represent the simultaneous application of several separable values.

+ Thus, the value premise "equality" may enter into norms for relationships between husband and wife, brother and brother, teacher and student, etc. On the other hand, the norm "a teacher must not show favoritism in grading" may in a particular instance involve the value of equality, honesty, humanitarianism and several others. Values, as standards (criteria) for establishing what should be regarded as desirable, provide the grounds for accepting or rejecting particular norm" [8].

+ Therefore, how are values, ethics and norms similar? It appears the intent of each is the promotion of a stable, safe and productive society. One can think of these as elements that, when they come together, help ensure that result. If normal behavior is working with a general sense of morality and the methods by which we make choices, then it can be expected that one has stability.

NEXT STEPS

+ It is still not fully decided on how far social norms overlap with values and how values overlap with ethics for a specific multi-cultural society, such as that in Canada.

+ If social norms prove to be appropriate substitutions of ethics, one could consider that they are advisable to prescribe implementation of norms in Autonomous Systems.

+ Specific algorithms, rules, and procedures for machine learning the social norms would be an ultimate challenge that would follow and deserve further development and investigation.

REFERENCES

[1] F. Alaieri, A. Vellino, Ethical Decision Making in Robots: Autonomy, Trust and Responsibility Auton- omy Trust and Responsibility, University of Ottawa.

[2] Royal Academy of Engineering, Safety and Ethics of Autonomous Systems, National Engineering Policy, Center, June 2020.

[3] T.J.M. Bench-Capon, Ethical Approaches and Autonomous Systems, Artificial Intelligence 281 103239, 2020.

[4] Stanford Encyclopedia of Philosophy, March 2011.

[5] Thomas Schelling, T. C. (1966), The strategy of conflict, New York: Oxford University Press.

[6] Thibaut, J.W. and Kelley, H.H. (1959), The social psychology of groups. John Wiley & Sons, New York.

[7] Ragnar Rommetveit (1955), Social Norms and Roles, Minneapolis: University of Minnesota Press.

[8] Puja Mondal, Essay on Values, Norms and Beliefs, https://www.coursehero.com/file/24476489/Meaning-and-Nature-of-Societydoc/.