2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
Madrid, Spain, October 1-5, 2018

The Future of Legal and Ethical Regulations for Autonomous Robotics

Huan Xu[1] and Joseph E. Borson[2]

***Abstract*— "Autonomous robotics" promise significant improvements across a host of different complex systems, which will need to be managed within regulatory frameworks to promote, at a minimum, device safety. Contrary to how they are often portrayed, however, these systems do not necessarily require fundamentally new approaches to engineering or regulatory challenges, i.e., the development of a novel "autonomy framework" applicable to different types of devices. Rather, because autonomous systems generally represent a progressive improvement of existing complex systems, preexisting regulatory scheme offer the best guidance for considering future regulation of autonomous elements. Moreover, the regulatory landscape differs considerably based on the type of device at issue (e.g., consumer electronics vis-à-vis medical devices). This paper argues that users and regulators must consider future autonomy regulations within the specific framework those devices currently inhabit, rather than focusing on a novel set of rules divorced from the preexisting context.**

I. INTRODUCTION

"Autonomous robotics," at least as popularly construed in the vein of self-driving cars, swarms of flying drones, or even military UAVs capable of deciding on their own to fire missiles, has been often been thought to be a fundamentally new type of technology, requiring fundamentally new types of engineering principles and regulatory approaches - an "autonomy" approach, as it were. This thought piece argues the opposite - autonomy is merely an evolutionary outgrowth of existing systems, and so existing systems engineering and regulatory approaches are the best ways of governing them. It does not make sense to speak of "autonomy" law, regulation, or even systems design as a new-in-kind category, or one that is generally applicable to different domains (e.g., aviation, automobiles, consumer products, etc.). Rather, existing frameworks - in particular, existing means of regulating complex systems (with their commensurate incorporation of different values and risk tolerances) are the best means to regulate autonomous systems.

There is not a universal definition of an autonomous system. Autonomy, however, at its core is merely the process of designing or regulating systems that make choices based not on contemporaneous, explicit, human intervention, but on pre-conceived decision rules. A self-driving car that determines how best to navigate a route based on an algorithm and sensor inputs is autonomous, but so is, in some sense, a stoplight that changes color based on pressure sensors, or even a medical device that delivers medicine based on a series of biological inputs. Accordingly, "autonomy" is not something new, rather, it is a measure of complexity that the law and the engineering fields have been confronting for many years.

Importantly, different fields have addressed autonomy challenges in different ways. Some fields, such as aviation, may prohibit autonomous devices altogether absent specific approval. Other fields, such as consumer electronics, may allow autonomous devices so long as they do not affirmatively interfere or cause problems. These approaches are often governed by the relative risk and benefits of the systems, or in some cases, by regulatory inertia or path dependency. Fields, for example, that are more heavily regulated or that require regulator pre-approval to begin with, like for certain medical devices, will follow these structural requirements with autonomous systems, just as they would with any system. Key, however, is that there is no single approach to regulation of autonomous systems - it is system and regulator specific. Accordingly, attempts to develop a unified theory of autonomous regulation, to the extent that such attempts are ongoing, are bound to fail, for the simple reason that they are necessarily unbound from the specific context of the systems they are trying to govern.

This work investigates what the legal and engineering communities believe are the current challenges in regulating autonomous systems, and posits that those challenges are not the true difficulties in legal and ethical regulation. This paper is structured as follows: Section II discusses the current challenges to regulation, Section III describes three application areas, and Section IV presents our conclusions.

II. CURRENT CHALLENGES

While there have been a number of articles in the news regarding the oncoming rush of autonomous systems into daily life, most have only posed concerns regarding what those changes could imply [1]. As the issue of regulation in autonomy is fairly new, the preliminary available literature in this area is somewhat limited. However, a few broad categories seem to emerge from existing work. These topics fall under the following categories: ethics, liability, and safety. We present a brief description of each of these issues, as well as reasons why the focus on these particular issues do not specifically address actual regulatory decisions.

A. Ethics and Society

1) Previous Work: A number of works focus on the issue of imposing morals on machines [2]. Anderson et al. [3] discusses the moral reasons why autonomous machines should function ethically. In the area of medical robotics, Stahl and Coeckelbergh [4] explore the implications of health

[1]Huan Xu is with Faculty of Aerospace Engineering, University of Maryland, College Park, MD, 20742, USA. mumu@umd.edu

[2]Joseph Borson is a graduate of Harvard Law School, Washington, DC, 20009, USA. jborson@jd13.law.harvard.edu

care robotics replacing human workers, both on the labor force as well as on the "humanization" and quality of care for patients. Wachter et al. [5] explore the need for transparent and accountable AI for robotics. Grinbaum et al. [6] claim that integration of robotics into society needs a new legal framework. They pose questions of robotics personhood, and when should machines should assume operator control. Finally, the authors in [7] discuss robotics and autonomy from a regulatory standpoint, examining existing legal structures and how adequate they are in encompassing the challenges that autonomy poses. They believe that there is a disconnect between traditional approaches to ethics of robotics and current innovation practices, and stress the need for embedded ethics within autonomous systems.

2) Actual Challenges: Much of the above literature sees autonomous systems as implicating fundamentally new ethical considerations. But in many cases, this overstates and obfuscates the challenge. Rather than representing a totally new system (the theorized self-aware robot from science fiction, for example), autonomous systems are a progressive outgrowth of existing complex systems. In other words, autonomy does not necessarily create a new-in-kind ethical challenge, but entails many of the same ethical challenges already present in modern system design: how to properly design a system that may take certain actions without direct human control, or one that may be so complex that it cannot be overseen (or perhaps even understood) by a single operator.

This work does not minimize these challenges. But focusing on the "end-game" questions of system personhood or full autonomy, as interesting as those questions may be in a theoretical sense, ignores the fact that most of the ethical questions likely to be seen in the near and medium term are going to be outgrowths of existing system regulation. Questions about autonomous medical devices, for example, would be informed by existing ethical standards governing medical regulations and techniques. Similarly, regulations of other types of vehicles will be governed by the degree of risk society is willing to accept for certain activities. If, for example, driving or flying is seen by the public as "risky," then it is likely that regulation of autonomous aspects of those activities will be governed more conservatively than activities that are seen as less risky (such as, for example, autonomous light systems for consumer purposes). That perceived risk calculus is further informed by the fact that autonomous systems are often seen by the public as "riskier" than other types of systems, even if the actual incidence of accidents would not warrant this concern [8]. In other words, while ethical considerations in different types of autonomous systems is important in the abstract, the proper lens of analysis is on the type of system involved, not necessarily on whether aspects of those systems are autonomous or not. In this sense, ethical considerations in different aspects of autonomous systems is likely be much more of an evolutionary outgrowth of existing considerations, rather than a revolutionary new consideration to be grafted on top of a field, simply because that field has autonomous elements.

B. Liability

1) Previous Work: Recent accidents involving self-driving cars has led to the question of insurance and responsibility [9]. Marchant and Lindor [10] explore the issue of liability in autonomous vehicle collisions, addressing the questions of who is liable, how risks determine liability, and what protections are available to manufacturers. Gurney [11] examines liability based on varying levels of driver control. For accidents involving service robots, Kelley et al. [12] suggest applying analogous laws governing domesticated animals.

2) Actual challenges: There are two fundamental questions with allocation liability for fault in any system: who bears the risk, and what should the consequences of fault be (or, said a different way, how much risk the regulator is willing to accept). As discussed above, much of the literature discusses the allocation problem. But the question, while important, can minimize the more important question of risk tolerance.

It clearly is relevant who bears the risk for a certain technology — in the self-driving car context, for example, if the manufacturer of a system is liable, that entity will behave differently than if the operator is liable. But focusing too much on who the liable actor ignores the fact that risk can be transferred, through, for example, insurance or other market forces [13]. Indeed, risk distribution — and thus liability distribution — is a common and relatively well understood phenomena in the legal literature, and is often, though not always, borne formally by the entity most able to pay, with costs passed down to other parties [14]. (For example, a manufacturer may bear the risk for a mis-designed device, but will pass the cost of bearing that risk to consumers in the form of higher prices).

Perhaps the more interesting question — and the starting point for the liability analysis — is what risk society is willing to bear (which will directly inform the degree of liability). And again, this will vary by field. In the context of autonomous systems in certain areas (such as critical infrastructure or perhaps transportation) the answer may be "not very much." In other systems, such as consumer contexts, the answer may be quite a bit more. (Which is why we accept certain failures with our iPhones, but not with our nuclear power plants.) And knowing that answer can inform the risk allocation — a risk that is not to be tolerated may be placed more directly on the entity creating the device, for a risk that can be more generally shared, liability, too, can be shared.

Similarly, the law already can distinguish between, broadly speaking, liability based on design defects and use defects. Such principles can be applied in the context of autonomous systems as well. Any "autonomous" system will still involve some operator input, even if just to set parameters or objectives, and it will, of course, involve design criteria. In the event of an accident, liability can be allocated based on the nature of the fault, just like it can be now. The fact that a system may have autonomous elements should not fundamentally change that calculation.

C. Safety

1) Previous Work: Safety of autonomous systems can be considered from both the engineering as well as regulatory standpoint. The design of future autonomous systems is in theory no different than the design of robotics systems currently on market. Indeed, Tesla motor vehicles, Boeing's 787 "Dreamliner," and DJI's numerous drone models all exhibit certain levels of "autonomy" that at times remove the human from the loop.

Generally speaking, the design of any complex system begins with both a top-down and bottom-up approach. From the top-down perspective, high-level requirements driven by stakeholders needs, customer utility, and engineering constraints form the basis for software and hardware architectures. From the bottom-up perspective, components, code, and models from legacy designs are modified and repurposed in order to reduce the idea of "redesigning the wheel." The reconciliation of these two approaches as well as the integration of different subsystems (e.g., thermal and power subsystems in an aircraft) then becomes the domain of certification officials in testing, evaluation, verification (i.e., is we build the system correctly), and validation (i.e., is the system the one we wanted to build). From a design perspective, [15] attempts to regulate autonomous systems via a systems engineering approach, while [16] adapts existing laws of armed conflict onto autonomous systems.

There are two fundamental types of regulations. The first, known as technology-based standards (or sometimes command-and-control, or proscriptive regulation), sets out a specific technological requirement that a complex system must include, for example, that all cars must include a catalytic converter [17]. This type of regulatory design governs inputs to a system. The second approach, known as performance-based standards, sets a specific performance standard that a system must satisfy, without specifically prescribing how that system is to achieve that outcome, e.g., mile-per-gallon requirements. These regulations govern a system's outputs.

Technology-based standards require precisely defined requirements to govern the system. For example, a regulation may specify specific types of sensors (or sensors with particular parameters), the inclusions of specific failure conditions (for example, that a device will shut down if the temperature crosses a specific threshold), or particular visibility or spacing requirements (in the context of highway signs). These regulations often have the benefit of being relatively straightforward to execute, at least in theory, and being relatively precisely defined. The costs, however, is that they can freeze technology development, if the state of the art has moved beyond older prescriptions, and can be difficult to adapt or apply in changed circumstances.

Performance-based standards, however, set the outcome that a system must satisfy, but are agnostic as to the method of achieving that outcome. In the environmental context, for example, a regulation may require that a certain compound be discharged at a certain percentage, such as 4.75 parts per million, but it would not set out specific technical requirements that an operator must satisfy in order to comply (as a technology-based standard would). Because these regulations are more focused on outcome than specific inputs, they can better allow for technological innovation and provide more flexibility given the specific needs of the system. The costs, however, are that these standards may be more difficult to define, validate, or assess, and may not be equally suited to all systems, particularly if the outcome is relatively ambiguous.

These two systems are not mutually exclusive: most regulatory systems, of course, use a mixture of these two types of regulations. FAA aircraft regulations, for example, specify both technology-based standards (such as requiring specific types of oil drain systems [18]) as well as performance-based standards (requiring that an aircraft can still reach a certain altitude "[i]f the critical engine fails at any time after the airplane reaches critical-engine failure speed [19]"). But determining which regulatory systems work best for which types of systems remains an important design challenge.

2) Actual Challenges: There is no single way of making an autonomous system "safe." Rather, safety considerations should be informed by the specific nature of the system: autonomous automobiles may be defined as "safe" based on different criteria than flying drones, as defined by the types of regulatory considerations described above. Figure 1 shows a breakdown of levels of automation as defined by the Society of Automotive Engineers [20]. Modern cars feature partial and conditional automation, while Google and Uber's self-driving cars fall somewhere in the category of 3, 4, and 5 (i.e., conditional, high, and full automation). Modern day aircraft, by these same standards, could be categorized as capable of high or full automation.

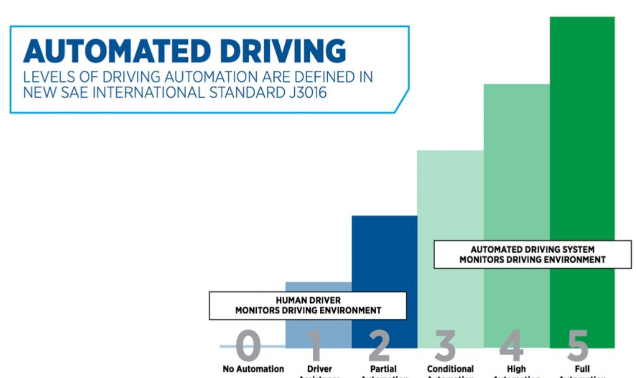

Fig. 1. The Society of Automotive Engineers (SAE)'s definition for the six levels of automation [20].

Accordingly, it does not make sense to think of a single set of "autonomy safety regulations," rather, safety regulations should be constructed based on the needs of a particular system, treating autonomous elements similar to how any complex element would be regulated (to the extent that a 'complex" element and an "autonomous" element are distinct).

III. APPLICATION AREAS

Approaches to regulation of autonomous systems will vary significantly depending on the application area. Certain fields may have different approaches to regulation, and the regulation of autonomous elements within those regulatory spaces will be informed by — and to a large extent, dependent on — the existing, field-specific regulatory frameworks.

Figure 2 depicts the various regulatory approaches for different systems. Consider first consumer electronics, such as a smart light or smart speaker system. To a large extent, these devices are relatively thinly regulated. While government entities may have the power to set standards governing product requirements [21], and in some cases to ban unsafe products, there is generally no requirement that these products be approved before going on the market. In other words, regulation generally applies if a product is demonstrated to be unsafe, rather than requiring an affirmative showing that the product is safe. Autonomous devices within this space would probably be treated the same way. In other words, to the extent that there are autonomous consumer devices, it is likely that regulatory efforts will focus on ensuring that unsafe devices will be taken off the market, rather than setting specific standards to ensure, pre-market, that the devices are themselves safe. Furthermore, it is unlikely that regulation in this sphere would focus on whether the products are effective (however that is defined), so long as they are not shown to be unsafe. Regulation of autonomous devices in this space, therefore, would likely be relatively light.

Self-driving cars would likely follow existing automotive regulations. Here, regulators may set specific safety standards that vehicles within a particular jurisdiction must comply with [22]. Manufacturers would generally have to certify that their vehicles comply with these standards before the vehicles may be sold, but the vehicles themselves would not need to be pre-approved by a government agency before going on the market [23]. For these types of products, system designers and regulators would likely have to work closely to define system requirements and translate those requirements into specific, workable standards (and there has indeed been already been discussion about the development of those standards) [24], [25]. Accordingly, existing automotive standards are a good model for autonomous vehicle standards.

Finally, consider medical devices. Here, these devices are subject to a much more stringent regime, in many cases requiring affirmative pre-market approval by a government entity, where the device manufacturer must demonstrate the safety and effectiveness of the device before it goes on the market [26]. These reviews are subject to rigorous ex ante standards, and autonomous medical devices would likely be modeled after existing regulatory standards.

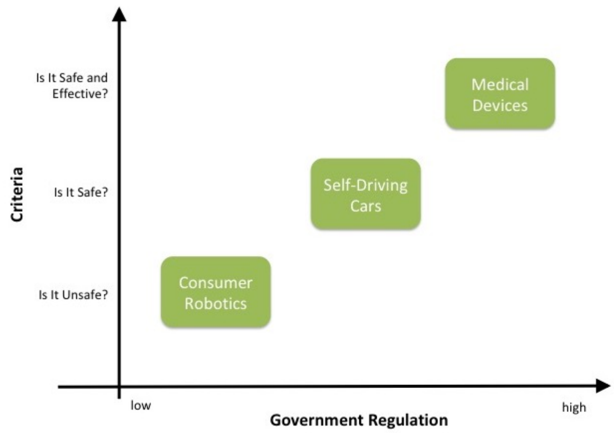

Fig. 2. Conceptualized regulatory approaches for different types of systems.

IV. CONCLUSIONS

In short, there is no single set of regulation for autonomous devices, rather, these devices will be regulated similar to others in their field, governed by similar processes and similar standard-setters. Rather than focusing on autonomy regulation writ large, it is likely more useful to focus on how autonomy plays out in the context of a specific type of field, and there, such regulation will be evolutionary, rather than revolutionary, relative to existing regulatory frameworks.

REFERENCES

[1] A. Bertolini, "Robots as products: the case for a realistic analysis of robotic applications and liability rules," *Law, Innovation and Technology*, vol. 5, no. 2, pp. 214–247, 2013.
[2] J. J. Bryson, "Patiency is not a virtue: Ai and the design of ethical systems," in *Ethical and Moral Considerations in Non-Human Agents, 2016 AAAI Spring Symposium Series*, 2016.
[3] M. Anderson and S. L. Anderson, *Machine ethics*. Cambridge University Press, 2011.
[4] B. C. Stahl and M. Coeckelbergh, "Ethics of healthcare robotics: Towards responsible research and innovation," *Robotics and Autonomous Systems*, vol. 86, pp. 152–161, 2016.
[5] S. Wachter, B. Mittelstadt, and L. Floridi, "Transparent, explainable, and accountable ai for robotics," 2017.
[6] A. Grinbaum, R. Chatila, L. Devillers, J.-G. Ganascia, C. Tessier, and M. Dauchet, "Cerna mission and context," 2017.
[7] R. Leenes, E. Palmerini, B.-J. Koops, A. Bertolini, P. Salvini, and F. Lucivero, "Regulatory challenges of robotics: some guidelines for addressing legal and ethical issues," *Law, Innovation and Technology*, vol. 9, no. 1, pp. 1–44, 2017.
[8] B. Schoettle and M. Sivak, "Motorists' preferences for different levels of vehicle automation," 2015.
[9] P. M. Asaro, "The liability problem for autonomous artificial agents," in *Ethical and Moral Considerations in Non-Human Agents, 2016 AAAI Spring Symposium Series*, 2016.
[10] G. E. Marchant and R. A. Lindor, "The coming collision between autonomous vehicles and the liability system," *Santa Clara L. Rev.*, vol. 52, p. 1321, 2012.
[11] J. K. Gurney, "Sue my car not me: Products liability and accidents involving autonomous vehicles," 2013.
[12] R. Kelley, E. Schaerer, M. Gomez, and M. Nicolescu, "Liability in robotics: an international perspective on robots as animals," *Advanced Robotics*, vol. 24, no. 13, pp. 1861–1871, 2010.
[13] R. H. Coase, "The problem of social cost," *The journal of Law and Economics*, vol. 56, no. 4, pp. 837–877, 2013.
[14] G. Calabresi, "Some thoughts on risk distribution and the law of torts," *The Yale Law Journal*, vol. 70, no. 4, pp. 499–553, 1961.
[15] D. P. Britton, "A systems engineering approach to regulating autonomous systems," Ph.D. dissertation, Duke University, 2017.
[16] K. Anderson, D. Reisner, and M. Waxman, "Adapting the law of armed conflict to autonomous weapon systems," *International Law Studies*, vol. 90, no. 1, p. 3, 2014.

[17] C. Coglianese, J. Nash, and T. Olmstead, “Performance-based regulation: Prospects and limitations in health, safety, and environmental protection,” *Administrative Law Review*, pp. 705–729, 2003.
[18] U. S. Government, “Code of federal regulations: Title 41, section 121.241,” 2017.
[19] ——, “Code of federal regulations: Title 41, section 121.177,” 2017.
[20] S. O.-R. A. V. S. Committee *et al.*, “Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems,” *SAE International*, 2014.
[21] U. S. Government, “Consumer product safety act, title 15, section 2051-2089,” 1972.
[22] ——, “Federal motor vehicles safety standards,” 1967.
[23] ——, “Code of federal regulations: Title 49, section 30115.”
[24] “Automated vehicles for safety,” September 2017. [Online]. Available: https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety
[25] T. Pearl, “Fast & furious: The misregulation of driverless cars,” 2016.
[26] U. S. Government, “Code of federal regulations: Title 21, part 814.”