Journal of Ethics and Legal Technologies – Volume 4(2) – November 2022

MaxSAT meets Alexy: A computational approach to solve trolley problem-like scenarios in the legal sphere

Samuela Marchiori

Faculty of Technology, Policy and Management
Delft University of Technology
s.marchiori@tudelft.nl

Abstract: In this paper, we investigate whether the computational approach for the ethical resolution of the trolley problem proposed by Sommaggio and Marchiori (2020) may be applied to trolley problem-like scenarios that transcend the ethical dimension. Specifically, we draw a parallelism between colliding ethical and legal principles, as well as between MaxSAT and Alexy's principle theory. We then apply the resolution strategy proposed by the authors to the legal domain, and present the potential benefits and shortcomings of a computational approach to legal reasoning.

Keywords: trolley problem; computational logic; legal reasoning; principle theory

1. Introduction

Since the mid-2010s, the trolley problem (TP) thought experiment has become prevalent in the discourse surrounding autonomous vehicles (AVs), i.e., vehicles with the ability to operate without the need for human intervention. Perhaps due to the bioethical roots of this thought experiment, the ethics of technology literature has dominated the field by extensively investigating the ethical concerns raised by the presence of AVs in trolley problem-like scenarios.

Interestingly, certain authors have transcended the merely philosophical sphere and approached the issue from an interdisciplinary perspective, e.g., by combining philosophy with cognitive sciences (Greene et al., 2009; Rai and Holyoak, 2010; Ahlenius and Tännsjö, 2012), psychology (Navarrete et al., 2012; Cao et al., 2017; Dupoux and Jacob, 2007), economics (Lanteri et al., 2008), and game theory (Hoffman et al., 2016; Jenkins et al., 2022). In this paper, we will investigate the computational approach for the resolution of the trolley problem proposed by Sommaggio and Marchiori (2020), which combines elements of moral and legal philosophy and computer science.

Specifically, we will investigate the extent to which, if at all, this approach can be applied to trolley problem-like scenarios that transcend the ethical dimension. We will argue that the computational approach proposed by the authors can be fruitfully applied to instances in which a trolley problem scenario presents itself as a collision between legal principles, instead of ethical ones, as it is traditionally the case. Additionally, we will put forth an example of how a computational approach for the resolution of legal disputes involving fundamental legal principles may take place, by applying this approach to examples proposed by Alexy.

2. Trolley problem and MaxSAT

Originally formulated by Philippa Foot (1967), the trolley problem dilemma consists of the following scenario. An out-of-control railway trolley is speeding along the railway tracks where five workers are working. Close to the tracks is a lever, which can be operated to divert the trolley onto another track, where one worker is working. At this point, one is presented with a dilemmatic choice: one has to decide whether to intervene and pull the lever, or not intervene and let the trolley continue on its original path. The former choice would result in the trolley being diverted onto the second track, and the worker on the second track being killed, while the latter would result in the trolley continuing its original course toward the five workers on the first track, killing them.

The advent of AVs has renewed interest in this thought experiment. In fact, the presence of vehicles capable of operating without the so-called human-in-the-loop (Zanzotto, 2019) opens up to scenarios in which machines may have to make life-or-death decisions comparable to those raised by the original trolley problem. Examples of TP-like scenarios include AVs experiencing brake failure and having to decide whether to prioritise the safety of their passengers or the safety of individuals travelling in nearby vehicles and pedestrians, as well as the case of pedestrians jaywalking, thus requiring AVs to decide whether to hit the pedestrians, or avoid them, while possibly endangering the safety of their passengers as a result. The possibility of TP-like scenarios arising in the context of AVs leads to two main issues.

The first issue concerns the necessity to instruct the AV as to how it should act in the context of a TP-like scenario before such a scenario has occurred. Indeed, TP-like scenarios are often characterised by their abruptness and require a swift (re)action. It follows that it would not be feasible to expect the AV to inform a human supervisor of an imminent TP-like scenario and wait for an *ad-hoc* human instruction. Instead, such a human supervisor should instruct the AV *ex ante facto*.

The second issue relates to different reasoning patterns between humans and machines. That is to say, it may not always be the case that a specific type of human reasoning may be faithfully translated into a language that the AV will also be able to comprehend, without losing nuance. For instance, it is worth noting that TP-like scenarios have been extensively analysed by applied psychologists and cognitive scientists (Greene, 2016; Petrinovich et al., 1993; Di Nucci, 2013; Bauman et al., 2014). Such studies have shown that, despite people having strong intuitions about how the trolley problem should be solved, they are oftentimes unable to express the reasoning behind their intuitions to a satisfactory degree. In this sense, it is not enough to know how to solve the trolley problem, if one cannot explain it to the AV in such a way that there is a high degree of likelihood that the AV's decision will mimic that of the person who instructed it.

Within this context, Sommaggio and Marchiori (2020) propose a computational approach for the resolution of TP-like scenarios. Specifically, the authors argue that one may consider the trolley problem in terms of a problem of maximum satisfiability, commonly known in computational complexity theory as MaxSAT (Maximum Satisfiability Problem). That is to say, instead of trying to find a one-right-answer to the trolley problem, one should rather consider it in terms of a multi-faceted issue with several variables, attribute a weight to each variable to determine one's preferences, and identify the model which, among all possible ones, maximises satisfaction.

As a way of example, one may consider the number of passengers traveling in the AV, the presence and number of pedestrians, the conditions of the road, and the integrity of the AV in terms of variables. Such variables will then be assigned positive, neutral, or negative weights depending on whether one wants to seek or avoid specific scenarios. For instance, one would reasonably assign a negative value to the loss of human life, while a positive value to the safety of passengers. Additionally, such values will range depending on their perceived salience. In this sense, it is reasonable to assume that the loss of a human life will be assigned a far lower value compared to damage to the AV, e.g., the former may be assigned -500, while the latter may be assigned -5. Moreover, this means that the choice as to whether one should or should not operate the lever may also be assigned a positive, neutral, or negative weight depending on whether the decision to intervene itself is considered value-laden, e.g., whenever one subscribes to the difference between "killing someone" and "letting someone die" (Thomson, 1976; Thomson, 1985; Unger, 1992; Unger 1996; Kamm, 1991).

According to this computational approach, one would represent the original trolley problem scenario in terms of the choice as to whether or not to operate the lever. In its more basic representation, the problem would include two variables, i.e., operating vs not operating the lever, and causing vs not causing a person's death. Let us represent them as follows.

L = the lever has been pulled
¬L = the lever has not been pulled

P_n = person *n* is alive
$\neg P_n$ = person *n* is dead

Subsequently, one would attach a cost, a weight, to each proposition (L, P_n, ¬L, $\neg P_n$), which would represent the preference for a certain event to (not) occur. By way of example, let us arbitrarily assign the value of -250 to the choice of operating the lever, and the value of -500 to the loss of a human life.[1]

L = -250
¬L = 0
$P_n = 0$
$\neg P_n = -500$

[1] The purpose of this example is to present a basic overview of how one may solve the trolley problem using MaxSAT. However, a real-case application of MaxSAT to TP-like scenarios would require a significantly more detailed account of all the variables that are worth being considered. For example, this will reasonable not be limited to attributing negative value to the loss of a human life, but will also include the attribution of a positive value in relation to the preservation of a human life.

In this sense, the original formulation of the trolley problem could be exemplified as a choice between L (which would imply $\neg P_1$), and $\neg L$ (which would imply $\neg P_2 \wedge \neg P_3 \wedge \neg P_4 \wedge \neg P_5 \wedge \neg P_6$). This would lead to two models.

$$\text{M1) } L \wedge \neg P_1 \wedge P_2 \wedge P_3 \wedge P_4 \wedge P_5 \wedge P_6$$

$$\text{M2) } \neg L \wedge P_1 \wedge \neg P_2 \wedge \neg P_3 \wedge \neg P_4 \wedge \neg P_5 \wedge \neg P_6$$

Let us calculate the cost of each model.

$$\text{M1) } -250 + (-500) + 0 + 0 + 0 + 0 + 0 = -750$$

$$\text{M2) } 0 + 0 + (-500) + (-500) + (-500) + (-500) + (-500) = -2500$$

From the determination of the costs associated to both models, it emerges that the first model would maximise condition satisfaction and should be preferred. This means that, if one were to use MaxSAT to face the trolley problem, one would opt for pulling the lever and diverting the trolley onto the second track, which would result in the death of one worker.[2]

3. MaxSAT meets Alexy

We argue that the key features of the computational approach proposed by Sommaggio and Marchiori (2020) may allow it to be fruitfully applied beyond the ethical dimension. That is to say, this approach may not only be useful to face TP-like scenarios where different moral values are colliding,

---

[2] A clarification is in order. Employing MaxSAT for the resolution of the trolley problem may lead to significant ethical concerns when one considers that the computational approach on which MaxSAT is built strongly evokes Bentham's felicific calculus (Bentham, 1789/2009). It follows that, while attributing a negative value to the choice of intervening may in practice preserve a semblance of deontological ethics (Kant, 1785/1993; Kamm, 2007), to the extent that one may assign values that are virtually insurmountable, this would likely not be enough to surpass the bias in favour of a utilitarian approach". That is to say, if one were to assign a value of -1000000 to the choice of pulling the lever, this might make it highly unlikely for a scenario to present itself where pulling the lever would be the choice that would maximise satisfaction. Nevertheless, while this model may seem capable of preserving some semblance of deontological ethics on the surface, it remains utilitarian at its core (Mill, 1985; Kamm, 2013), in that it does not provide the choice of abstaining from judgement.
It is also worth mentioning that, in previous work, Sommaggio and Marchiori (2018, 2020) analysed traditional and contemporary utilitarian and deontological approaches to the trolley problem, both from a predominantly philosophical perspective and in combination with studies in applied psychology and cognitive sciences, and ultimately favoured the utilitarian approach. In this sense, pursuing a decidedly utilitarian approach to the trolley problem was a deliberate choice from part of the authors.

thereby providing guidelines for the adoption of one or more value-laden choices over some others, but it may also prove helpful within the legal dimension, whenever legal principles are colliding and need balancing. For the purpose of this investigation, we will rely on Robert Alexy's principle theory, which we will use as a starting point to illustrate how MaxSAT may be a useful tool to guide the resolution of legal controversies.

Let us first clarify the distinction between legal rules and legal principles. According to Alexy, while legal rules can be based on multiple legal principles, legal principles - like prime numbers - cannot be inferred from anything other than themselves or the very idea of principle. Thus, legal principles can hardly be justified in terms other than tautologies and general statements of the form "because", e.g., why is dignity important? *Because.*[3]

Briefly, three core ideas lay the foundations for Alexy's principle theory, i.e., the optimisation thesis, the collision law, and the balancing law. The optimisation thesis holds that principles are "norms commanding that something be realized to the highest degree that is actually and legally possible" (Alexy, 2000, 295). In this sense, principles are considered as "optimization commands", whose degree of fulfilment depends on legal possibilities (Alexy, 2000; Alexy 1985/2010; Recht, 1995). The collision law holds that "the conditions under which one principle takes priority over another constitute the operative facts of a rule giving legal effect to the principle deemed prior" (Alexy, 2000, 297; Alexy, 1985/2010). Lastly, the balancing law states that "the more intensive the interference in one principle, the more important the realization of the other principle" (Alexy, 1985/2010, 146; Alexy, 2000).

The first one is particularly relevant for the investigation at hand. In fact, in his theory, Alexy does not merely argue that legal principles require balancing, as many legal scholars do, but goes on to identify such a balancing with the application of the principle of proportionality, which "necessarily involves optimization" (Alexy, 2003; Tschentscher, 2012), thus considering legal principles as opmitisation commands (Alexy, 2000; Alexy, 2014; Alonso, 2016; Greer, 2004). In this sense, Alexy's principle theory seems to closely resemble the kind of approach one would have to follow to implement MaxSAT in a legal setting (Sieckmann, 2010; Menéndez and Eriksen, 2006).

That is to say, according to Alexy's theory, whenever legal principles come into the picture, a duty to optimise them always follows. This means that, if there are two ways to promote the same interest, the alternative should be chosen which causes minor interference with individual rights, while still promoting that interest. This seem to be in rather close alignment with the

[3] It is worth noting that Alexy's position is not without its criticisms. In this sense, see Ramião (2018), Martínez-Zorrilla (2018), and La Torre (2006).

concept of maximum satisfiability. Let us test the applicability of MaxSAT to problems formulated according to the Alexyan principles theory by considering one of Alexy's own examples, and try to represent this issue in terms of a MaxSAT calculation. The example states as follows.

> Let us assume a measure M that encroaches on the freedom of trade, occupation, or profession (P_1) in order to promote consumer protection (P_2) but which is not appropriate to promoting P_2 in any way whatever. It is possible to abandon M at no cost to P_2, consumer protection. The optimization of P_1 and P_2 demands, then, that M not be used. This is exactly the content of the principle of appropriateness. The principle of necessity says that a measure M_1 is prohibited in respect of P_1 and P_2 if there is an alternative measure M_2 that promotes P_2 approximately as well as M_1 but encroaches less intensively on P_1. Let us assume that P_2 stands, again, for consumer protection, in particular, for the consumers' protection against buying products that they do not in fact want. Let us also assume that M_1 is an absolute prohibition of goods that look like chocolate but are not chocolate. M_2 stands in this case for the obligation clearly to designate the nature of the goods. This obligation, namely (M_2), obviously encroaches less intensively on the freedom of trade, occupation, or profession (P_1) than would an absolute prohibition (M_1), and it serves consumer protection more or less equally well; therefore, the absolute prohibition (M_1) is prohibited in relation to P_1 and P_2 as an unnecessary means [...] (Alexy, 2000, 298).

For reasons of convenience, we will divide the different instances of P_1 and P_2 as presented in the two scenarios[4] as F_1 and P_1, F_2 and P_2, respectively, and represent the problem in three stages. Firstly, let us present the data as follows.

M_1 = measure 1 is applied
$\neg M_1$ = measure 1 is not applied
M_2 = measure 1 is applied
$\neg M_2$ = measure 1 is not applied
F_1 = freedom of trade is preserved
$\neg F_1$ = freedom of trade is heavily restricted
F_2 = freedom of trade is preserved
$\neg F_2$ = freedom of trade is partially restricted
P_1 = consumer protection is guaranteed
$\neg P_1$ = consumer protection is not guaranteed
P_2 = consumer protection is guaranteed
$\neg P_2$ = consumer protection is not guaranteed

[4] By this we mean the first scenario, which applies the measure M_1, and the second one, which applies M_2.

In this sense, M_1 implies $\neg F_1 \wedge P_1$, $\neg M_1$ implies $\neg F_1 \wedge \neg P_1$, M_2 implies $\neg F_2 \wedge P_2$, $\neg M_2$ implies $\neg F_2 \wedge \neg P_2$, and $M_1 \wedge M_2$ implies M_2. Let us also arbitrarily assign some weights to our constraints.

$M_1 = 0$	$\neg M_1 = 0$
$M_2 = 0$	$\neg M_2 = 0$
$F_1 = 0$	$\neg F_1 = -100$
$F_2 = 0$	$\neg F_2 = -50$
$P_1 = 0$	$\neg P_1 = 0$
$P_2 = 0$	$\neg P_2 = 0$

Let us now consider the first scenario and represent it in terms of MaxSAT:

> Let us assume a measure [M_1] that encroaches on the freedom of trade, occupation, or profession [F_1] in order to promote consumer protection [P_1] but which is not appropriate to promoting [P_1] in any way whatever. It is possible to abandon [M_1] at no cost to [P_1], consumer protection. The optimization of [F_1] and [P_1] demands, then, that [M_1] not be used (Alexy, 2000, 298).

$$\text{M1a) } M_1 \wedge \neg F_1 \wedge P_1$$
$$0 \wedge -100 \wedge 0$$
$$= -100$$

$$\text{M1b) } \neg M_1 \wedge F_1 \wedge \neg P_1$$
$$0 \wedge 0 \wedge 0$$
$$= 0$$

Let us now consider the second scenario, and represent it in terms of MaxSAT:

> Let us assume that P_2 stands, again, for consumer protection [...] M_2 stands in this case for the obligation clearly to designate the nature of the goods. This obligation, namely (M_2), obviously encroaches less intensively on the freedom of trade, occupation, or profession [F_1] than would an absolute prohibition (M_1), and it serves consumer protection more or less equally well (Alexy, 2000, 298).

$$\text{M2a) } M_2 \wedge \neg F_2 \wedge P_2$$
$$0 \wedge -50 \wedge 0$$
$$= -50$$

$$\text{M2b) } \neg M_2 \wedge F_2 \wedge \neg P_2$$
$$0 \wedge 0 \wedge 0$$
$$= 0$$

Once again, MaxSAT confirms that M_2 is preferable to M_1. Let us now consider the third and last scenario, and represent it in terms of MaxSAT:

> The principle of necessity says that a measure M_1 is prohibited in respect of P_1 and P_2 if there is an alternative measure M_2 that promotes P_2 approximately as well as M_1 but encroaches less intensively on P_1 (Alexy, 2000, 298).

$$\text{M3) } (M_1 \wedge \neg F_1 \wedge P_1) \vee (M_2 \wedge \neg F_2 \wedge P_2)$$
$$M_2 \wedge \neg F_2 \wedge P_2$$
$$0 \wedge -50 \wedge 0$$
$$= -50$$

This would seem to suggest that MaxSAT may be fruitfully applied to approach, and potentially solve, legal issues that may be reminiscent of a TP-like scenario, but do not necessarily encompass (moral) value judgments. However, this computational approach presents several limitations, among which two are particularly significant.

- First and foremost, how should weights be attributed to variables? Who should be in charge of such a determination? On which grounds should this evaluation take place? One may argue that such an assessment should be carried out by the scientific community of reference. While this may seem like a reasonable assumption, it is essential that such a determination represent an accurate reflection of the scientific community at large, and be diverse and inclusive enough to both account for different perspectives and to prevent individual biases from becoming (or rather, remaining) entrenched in the legal system.
- Second, a computational approach to legal controversies may be perceived as unfair, both by legal scholars and by lay people, to the extent that algorithmic decision-making will never be able to fully mimic human reasoning. As such, regardless of how close MaxSAT will be able to perform compared to human legal reasoning, it will never be good enough. A solution to this objection may be that of considering MaxSAT as a useful instrument to accompany traditional legal reasoning, without substituting it *in toto*. This may give legal scholars and practitioners the tools to determine the optimal course of action, while still deciding whether or not to follow MaxSAT's suggested model.

4. Conclusion

In this paper, we investigated whether the computational approach put forth by Sommaggio and Marchiori (2020) can be fruifully applied to legal disputes. To this end, we first illustrated the key features of such an approach, by highlighting how MaxSAT can provide guidelines as to how ethical principles should be balanced in trolley problem-like scenarios. Subsequently, we drew a parallelism between conflicting ethical principles and conflicting legal principles, and illustrated how MaxSAT may be compatible with Alexy's principle theory. We concluded by highlighting the main shortcomings related to the application of such an approach to legal controversies, i.e., the way in which the determination of the weights to be attributed to the variables take place, and the inherent limits of translating human legal reasoning in the form of algorithms.

Nevertheless, the potential benefits of such a computational approach deserve further attention. For instance, could the systematic application of this approach contribute to diminishing the degree of arbitrariness in the solution of legal disputes? Moreover, could it contribute to safeguarding the principle of legal certainty?

References

Alexy, Robert. *A Theory of Constitutional Rights.* Oxford: Oxford University Press, 1985/2010.

Alexy, Robert. "On the Structure of Legal Principles." *Ratio Juris* 13, no. 3 (2000): 294-304.

Alexy, Robert. "Constitutional Rights, Balancing, and Rationality." *Ratio Juris* 16, no. 2 (2003): 131-140.

Alexy, Robert. "Formal Principles: Some Replies to Critics." *International Journal of Constitutional Law* 12, no. 3 (2014): 511-524.

Alonso, Juan Pablo. "The Logical Structure of Principles in Alexy's Theory. A Critical Analysis." *Revus. Journal for Constitutional Theory and Philosophy of Law* 28 (2016): 53-61. https://doi.org/10.4000/revus.3533.

Ahlenius, Henrik, and Torbjörn Tännsjö. "Chinese and Westerners Respond Differently to the Trolley Dilemmas." *Journal of Cognition and Culture* 12, no. 3-4 (2012): 195-201.

Bauman, Christopher W., A. Peter McGraw, Daniel M. Bartels, and Caleb Warren. "Revisiting External Validity: Concerns about Trolley Problems and other Sacrificial Dilemmas in Moral Psychology." *Social and Personality Psychology Compass* 8, no. 9 (2014): 536-554.

Bentham, Jeremy. *An Introduction to the Principles of Morals and Legislation.* New York: Dover Publications Inc, 1789/2009.

Cao, Fei, Jiaxi Zhang, Lei Song, Shoupeng Wang, Danmin Miao, and Jiaxi Peng. "Framing Effect in the Trolley Problem and Footbridge Dilemma: Number of Saved Lives Matters." *Psychological Reports* 120, no 1 (2017): 88-101.

Di Nucci, Ezio. "Self-sacrifice and the Trolley Problem." *Philosophical Psychology* 26, no. 5 (2013): 662-672.

Dupoux, Emmanuel, and Pierre Jacob. "Universal Moral Grammar: A Critical Appraisal." *Trends in Cognitive Sciences* 11, no. 9 (2007): 373-378.

Foot, Philippa. "The Problem of Abortion and the Doctrine of Double Effect." *Oxford Review* 5 (1967): 5-15.

Greene, Joshua. "Solving the Trolley Problem." *A Companion to Experimental Philosophy* (2016): 175-178.

Greene, Joshua D., Fiery A. Cushman, Lisa E. Stewart, Kelly Lowenberg, Leigh E. Nystrom, and Jonathan D. Cohen. "Pushing Moral Buttons: The Interaction Between Personal Force and Intention in Moral Judgment." *Cognition* 111, no. 3 (2009): 364-371.

Greer, Steven. "'Balancing'" and the European Court of Human Rights: A Contribution to the Habermas-Alexy Debate." *The Cambridge Law Journal* 63, no. 2 (2004): 412-434.

Hoffman, Moshe, Erez Yoeli, and Carlos David Navarrete. "Game Theory and Morality." In *The Evolution of Morality,* 289-316. Cham: Springer, 2016. https://doi.org/10.1007/978-3-319-19671-8_14.

Jenkins, Ryan, David Cerny, and Thomas Hribek. *Autonomous Vehicle Ethics: The Trolley Problem and Beyond.* Oxford: Oxford University Press, 2022.

Kamm, Frances. "Harming Some to Save Others." *Philosophical Studies* 57, no. 3 (1989): 227-260.

Kamm, Frances. *Intricate Ethics: Rights, Responsibilities, and Permissible Harm.* Oxford: Oxford University Press, 2007.

Kamm, Frances. "Trolley Problem." *International Encyclopedia of Ethics* (2013): 1-6.

Kant, Immanuel. *Grounding for the Metaphysics of Morals.* Indianapolis, IN: Hackett, 1785/1993.

La Torre, Massimo. "Nine Critiques to Alexy's Theory of Fundamental Rights." *Arguing Fundamental Rights*, 53-68. Dordrecht: Springer, 2006.

Lanteri, Alessandro, Chiara Chelini, and Salvatore Rizzello. "An Experimental Investigation of Emotions and Reasoning in the Trolley Problem." *Journal of Business Ethics* 83, no. 4 (2008): 789-804. https://doi.org/10.1007/s10551-008-9665-8.

Martínez-Zorrilla, David. "Some Thoughts about the Limits of Alexy's Conception of Principles and Balancing." *Proportionality in Law*, 171-192. Cham: Springer, 2018.

Menéndez, Augustín J., and Erik O. Eriksen. *Arguing fundamental rights*. Dordrecht: Springer, 2006.

Mill, John Stuart. *Utilitarianism*. London: Longmans, Green and Company, 1895.

Navarrete, C. David, Melissa M. McDonald, Michael L. Mott, and Benjamin Asher. "Virtual Morality: Emotion and Action in a Simulated Three-dimensional 'Trolley Problem'". *Emotion* 12, no. 2 (2012): 364-370. https://doi.org/10.1037/a0025561.

Petrinovich, Lewis, Patricia O'Neill, and Matthew Jorgensen. "An Empirical Study of Moral Intuitions: Toward an Evolutionary Ethics." *Journal of Personality and Social Psychology* 64, no. 3 (1993): 467-478. https://doi.org/10.1037/0022-3514.64.3.467.

Rai, Tage S., and Keith J. Holyoak. "Moral Principles or Consumer Preferences? Alternative Framings of the Trolley Problem." *Cognitive Science* 34, no. 2 (2010): 311-321. https://doi.org/10.1111/j.1551-6709.2009.01088.x.

Ramião, Rúben. "Some Fundamental Problems Concerning Alexy's Notion of Legal Principles." *Proportionality in Law*, 157-170. Cham: Springer, 2018.

Sieckmann, Jan-Reinard. "Balancing, Optimisation, and Alexy's 'Weight Formula'". *Legal Reasoning. The Methods of Balancing* (2010): 101-118.

Sommaggio, Paolo, and Samuela Marchiori. "Break the Chains: A New Way to Consider Machines' Moral Problems." *Biolaw Journal*, 3 (2018): 241-257.

Sommaggio, Paolo, and Samuela Marchiori. "Moral Dilemmas in the AI Era: A New Approach." *Journal of Ethics and Legal Technologies* 2, no. 1 (2020): 89-102.

Thomson, Judith Jarvis. "Killing, Letting Die, and the Trolley Problem." *The Monist* 59, no. 2 (1976): 204-217.

Thomson, Judith Jarvis. "The Trolley Problem." *Yale Law Journal* 94, no. 6 (1985): 5.

Tschentscher, Axel. "Interpreting Fundamental Rights. Freedom versus Optimization.", *Debates in German Public Law* (2012): 43-56.

Unger, Peter. "Causing and Preventing Serious Harm." *Philosophical Studies* 65, no. 3 (1992): 227-255.

Unger, Peter. *Living High and Letting Die: Our Illusion of Innocence.* Oxford: Oxford University Press, 1996.

Zanzotto, Fabio Massimo. "Human-in-the-loop Artificial Intelligence." *Journal of Artificial Intelligence Research* 64 (2019): 243-252.