



Delft University of Technology

Document Version

Final published version

Citation (APA)

Copeland, S. M. (2024). Ethics as Generative Modelling. In *Model-Based REasoning, Abductive Cognition, Creativity: Inferences and Models in Science, Logic, Language, and Technology* (pp. 66-74). Article 5 (SAPERÉ Studies in Applied Philosophy, Epistemology and Rational Ethics). Springer. https://doi.org/10.1007/978-3-031-69300-7_5

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

In case the licence states "Dutch Copyright Act (Article 25fa)", this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.



Ethics as Generative Modelling

Samantha Copeland^(✉)

TU Delft, Delft, The Netherlands
S.M.Copeland@tudelft.nl

Abstract. In this chapter, Samantha Copeland explores the relationship between ethics practice and theory and recent work in modelling theory. Starting with model-based reasoning as theorized by Magnani and Nersessian, Copeland draws from recent work on normative modelling as well as recent work on participatory multi-modelling. The parallels reveal both descriptive commonalities as well as grounding normative advice for the practice of ethics in the contemporary world.

1 Introduction

Let us start at the beginning – why propose that ethics is a kind of generative modelling? In part, this paper follows a train of thought begun with a sentence from Lorenzo Magnani: “When looking for consequences of our moral actions and deliberations, the envisioning of a non-existing world as a means for judging a proposed action can be performed in a model-based way” (2006, p.220). In turn, that train intersects with another, on a track in constant development as I continue in my role as educator and ethicist. A further junction occurs with my work as a philosopher among modellers, at a technical university in a systems engineering faculty. This paper thus brings together considerations about the intersections between participatory modelling and social learning, and our expectations of ethical theory and practice. In it, I trace both the reasons why we might best understand (much of) ethics theory-building through the lens of model-based reasoning, and why utilizing the paradigms and insights of modelling practice can enhance our capacity to make an impact in ethical practice.

Neither I nor Magnani are the first to suggest that ethics and modelling are similar practices, or that modelling is useful in ethics. Nersessian (1999), for example, notes the use of model-based reasoning through analogical thinking and thought experiments, and both are long-standing tools in ethics. As I show below, there are many reasons to see ethics work, be it theory development, education or advice, as employing model-based reasoning or models themselves. Modelling theory has a potential role in supporting both ethics theory development and application: modelling is known to be useful for determining the potential consequences of actions before they are taken (i.e. Exploratory modelling), and they can reveal assumptions about relationships (cf. Epstein 2008; cited in Edmonds et al. 2019, 1.4). New developments in participatory, transdisciplinary modelling has led to new techniques for using models in the creation of boundary objects and shared understandings that reflects the goals of ethics practice. Thus, this paper further suggests that we ought to take up these new ideas in model-based reasoning into ethics

as guides for adopting more inclusive, pluralistic approaches to addressing increasingly complex ethical issues. In addition, in order for model-based reasoning to assist in this ethical context, multi-models must be our approach and generative ethics our aim. That is, ethics, I hope to show, is a *creative* enterprise, and modelling a tool through which assumptions can be laid bare, and the emergence of insights through deliberation made possible.

2 Is It Appropriate to See Ethics as a Kind of Modelling?

There are several familiar types of models used in ethics—I noted above thought experiments and analogies are common. Perhaps the most obvious ‘model’, however, is the Utopia: a utopia presents a complex ethical system in which the rules are followed and cause and effect relations are understood and thus predictable. Consider Kant’s Kingdom of Ends, in which all present are equally rational and equally respecting of each other’s rationality, thereby each fully expected to come to the same rational conclusions about the choices that ought to be made. I choose the Utopia to focus on as a kind of ethical model to point out its flaws—this is not the kind of model-based thinking I propose we utilize to theorize or to explain ethics. Indeed, one might even say that such a model presents a world in which ethics is no longer needed: why build a model to test hypotheses, when predictions are guaranteed? A model of a utopia is closer to a work of art or a toy: a finished product.

Roussos (2022) says something similar when he draws out how ethics and modelling in science are alike—unlike utopia, scientific models do not attempt to perfectly capture a precisely ordered ideal system. Rather, they perform a function and make assumptions to do so. Similarly, he argues, ethical theories abstract from reality in order to allow the testing of hypotheses about what the right action would be in a situation. The benefit of seeing this exercise through the lens of model-based reasoning, suggests Roussos, is that if we follow the wisdom of identifying the assumptions we have to make in modelling as well as being clear about our model’s function, as we would do in science, we can use ethical models also similarly to test our possible actions against their possible real-world consequences. Further, model-based reasoning offers the means to describe ethics as a pluralistic enterprise, something I develop further in the following section.

Insofar as ethics is a practice of theory development, it employs means for creating or enabling conceptual change. Nersessian and Magnani both argue clearly that model-based reasoning includes ampliative reasoning—that is, via modelling, new knowledge can emerge. Magnani describes manipulative abduction, the use of enaction to derive potential explanations for a phenomenon, or to derive phenomena that can offer new insight into the model or about its target. This ‘thinking through doing’ is an important practice in ethics, for instance when we employ thought experiments such as the trolley problem to identify conflicts between theory and intuition about what the right thing to do would be. Such exploratory models have frequently led to theoretical development and refinement in philosophical ethics.

Indeed, Roussos (2022) suggests that ethics is a kind of modelling because it is “an indirect strategy of inquiry” (p. 870). Thought experiments have a particular purpose, “and it is relative to that purpose that its results should be evaluated and made use of”

(p. 874). Like scientific models, they are necessarily incomplete but, by definition, useful. The examples Roussos provides of normative modelling include thought experiments, but also paradigm cases, from which principles might be derived (p.881). Models, that is, can be and are frequently used creatively to develop ethical theory.

Insofar as ethics is a pedagogical practice--offering the means for people to make good choices, for example--it should employ methods for generating insights and introducing new knowledge and meaning to its students. For instance, thought experiments or analogies frequently play a role in creating an experiential learning opportunity in the classroom, by triggering an insight in a student who is faced with the conflict between what they thought they knew and how they thought they ought to act in the situation presented. In the following section, I suggest that thinking of ethics as generative modelling has several positive implications for practice and theory in ethics.

3 Implications of Theorizing Ethics as Generative Modelling

As mentioned at the end of the last section, (not only pedagogical) ethics ought to have as one of its aims the production of insights in its students (i.e. Anyone seeking to learn about any ethics). This aim is very similar to the function of social learning that many modelling exercises have: “Sometimes our simulations are not about the observed world at all, but (at least substantially) designed to reflect an actor’s view of that world. Such simulations can be developed collectively by a group of people. In this case, the model can act as a mediator between members of the group and can result in a shared understanding of the world (or at least a clear idea where the members agree and differ).” (Edmonds et al. 2019, 8.1). I will first explore this idea of models as mediators toward shared understanding in ethics.

I have already indicated two ways that models can be mediators in ethics. First, insofar as they can be “boundary objects”, around which deliberation can occur and shared understanding obtained. Second, they can act as tools with which we can test various theoretical formulations for their validity and/or test proposals for practical actions for their soundness. Each use suggests something different about ethics itself; first, if boundary objects are useful in ethics, this is because of the plurality of ethical positions and potentially valid principles that may yet conflict. Second, if models can be used to test potential theory and practice in ethics, then ethics itself is a generative exercise: on the one hand, using models in this way both generates insights and can (and is intended to) lead to conceptual change, and on the other hand, there is good reason to prefer generative over degenerative processes in ethics, as the following comparisons with model-based thinking will show.

3.1 Models as Boundary Objects in Ethics

As Roussos notes (p. 873), in model-based thinking, multiple models are frequently used to illustrate various aspects of a problem situation, each one focussing on a different aspect. Taken together, multi-models enrich understanding of how things relate to each other; their purpose is not to produce a unified, stable model but rather to explicitly recognize the diversity of relevant perspectives on a problem as part of the complex

situation rather than something to be reduced (Nikolic et al. 2019, p. 3). Similarly, says Roussos, in ethics we look to understand a problem from diverse perspectives, without reducing it.

If we draw from multi-modelling theory, ‘participatory multi-modelling’ seems the closest in process and aim to ethics pedagogy and practice (*cf* Cuppen et al. 2020): “The structured, systematic, and collective activity of conceptualising, constructing, and using models by a group of stakeholders [...] provides unique insights to the stakeholders and ample opportunities for social learning.” (p. 903). In such a context, the process of modelling is a collaborative process, by which agreements are made about priorities, interpretations, rules and goals. The deliberations themselves offer insights into the perspectives of those involved in the model-making process.

Thus, if we think of multi-models as an outcome, it offers a way of seeing how diversity can be useful and informative in ethics, suggesting that pluralism is appropriate in ethics as well, or at least that it does not necessarily lead to incommensurability nor relativism. Multi-models consist of interactive models, that may require interpretation or adjustment to relate to one another, but overall find overlaps in purpose and content that are themselves informative, whether they are supportive or conflicting relations. In dilemma exercises, ethicists and their students seek to find compromise by analysing the case and the outcomes of possible actions; in these cases, the goal is a shared understanding that leads to a common interpretation of what is important about the case and shared decision about what the options are. Multi-modelling as a framework further opens up the potential for exploring multiple aspects and options without committing to consensus as the goal. Ethics deals with dilemmas and what many call “wicked problems”: “when dealing with wicked problems, the *process* of modelling is more important than the results of a model” (Cuppen et al. 2020, p.902). The focus thus becomes understanding how the interaction between models may affect them all—in ethical terms, how one’s perspective when enacted can have impact on another’s world.

The idea of models as boundary objects, however, offers an additional way of understanding pluralism in ethics practice, as a means toward consensus in understanding, if not definitive or predictable agreements about what is the right thing to do. In a participatory, multi-modelling context, a number of boundary objects will arise during negotiations between parties over what is important to model and how. Cuppen et al. (2020) describe this as a “boundary object ecology”: “an interacting group of boundary objects that interact and co-evolve with one another within the context of a dynamic participatory process” (p. 904). That is, boundary objects are meant to be shared, and thus they are negotiable and emerge from the relations between participants; like those relations, they are thus dynamic. Model-based reasoning thus has something to offer to ethics, a way of thinking about how the principles and frameworks we offer can be seen as multi-models that interact in ethics practice. Rather than attempting to consolidate all ethics under a single framework or in a single (utopian) model, the purpose of ethics can be seen as attending to those very interactions—how are they being negotiated, and what insights do they provide?

3.2 Models as Ethical Tools

A focus on process and interactions rather than outcome lies in contrast to another idea found in and about ethics, that principles and frameworks offer generalizable and accessible advice—beforehand—about what decisions and actions are the right ones. That is, good ethical theory will lead to good advice about what counts as good ethical practice. It is well known, however, that most ethical principles require a degree of interpretation before they can be applied to a specific situation, so it is not unusual for them to fall short of offering direct or consistent advice. A model, however, can provide the means for testing out possible actions and decisions, to see what would happen, and thus can be thought of as similar to other ethical tools—like thought experiments, analogies and others I have already mentioned. An ideal model, under the interpretation of ethical principles as guides, will offer an accurate interpretation of what would happen, if one decision were made instead of another.

But as we have seen so far, model-based thinking shows us that the ideal, stable and universalizable model is not the goal of models, modellers or model-based thinking. Nersessian relates model-based thinking to reasoning through modelling (2019, p.20) and Magnani coins the idea of thinking through doing (2006, p. 217) to describe it. If we think of ethics in terms of modelling, then, we can take up the ideal that it is the enaction of the principles through model-building that offers insights, rather than (only or merely) testing one outcome against another, according to the principles as criteria. Like the exercise of reflective equilibrium, that is, model-building in ethics would consist in an iterative consideration of how the model reflects the world, both the world we want and the world we live in. In the next sub-section, I show how this leads to a generative idea of ethics theory-building, pedagogy and practice—it is not about testing our principles against the world only, but also about generating new concepts and possibilities for action toward a more ethical world.

3.3 In Sum, and a Defence of Generative Ethics

Magnani (2006 p.213) draws from Feyerabend in a way that shows how model-based reasoning and dilemma-based reasoning can differ significantly. Via “counterinduction”, Feyerabend proposes a more creative and deconstructive approach to scientific investigation than using experiments to prove hypotheses (in)correct. As Magnani describes it, the recommendation is to draw out inconsistencies in current theory, but more importantly, to introduce new rival views. Like in dilemma-based approaches, experiments designed to prove hypotheses aim to eliminate one view in favour of another; Feyerabend’s creative approach to science could be applied to ethics, insofar as the point of deliberation could be to draw out new possibilities for action rather than diminish the ones available.

As suggested above, this means an anti-utopian approach to modelling in ethics. It is also in line with current theory in model-based reasoning. From philosophy, Nersessian shows that modelling can lead to conceptual change, allowing insights into how things relate to one another by embodying those relationships (1999, p. 6). Magnani (2006, p. 209, 213) elaborates on this as manipulative abduction, utilizing objects (e.g. Boundary objects, models) to draw out implications of principles and constraints—i.e.

To find explanations for what one sees happening in the world, and thereby create new knowledge.

As Magnani (2006) points out, even when we use models to deliberate via comparing results of different actions, as “a sort of selection or creation of principles and to their application to concrete cases” (p.206), frequently we lack the principles we need or find those we have to be incomplete or inappropriate. In such cases, creativity is required. Multi-model and participatory contexts encourage not competition but sharing, and toward a plurality rather than a singular outcome. Similarly, generative ethics focusses on how better understanding of complex situations (i.e. Wicked problems) can lead to the emergence of not only understanding but novel (i.e. Not predicted) outcomes. That is, reflective equilibrium and other methods in ethics tend toward utilizing common principles toward solutions, whereas model-based reasoning in ethics would and does instead create the conditions for new perspectives and shared understandings to emerge. The former, like utopia, moves toward a singular and universal set of principles or framework, and has the aim of prediction. The latter, like modelling in scientific practice, has a function and creates boundary objects over which deliberation can take place, with exploration and creativity as its aim. The division thus produced is not between ethical theory and its application, but rather between a degenerative and generative ethics.

4 Counterarguments Considered

Several counterarguments present themselves to the approach to ethics as generative modelling described so far. Here we entertain two, one that addresses the concern about changing from an outcome-based to a procedural approach to what is ethical, and a second that addresses the assumptions made in this paper about the nature of ethics.

4.1 Exchanging the Goal of Ethics Practice and Theory from an Ethical Outcome to Having Instead the Goal of a Generative Process Means that We Cannot Guarantee the Outcome of Our Process Will Be Ethical

One implication of this is that any process, to be ethical, will have to be guided by individuals with ethical expertise in order to lead it toward an ethical outcome. Thus, it seems we have not actually exchanged the goal of ethics but only reinforced it, as ‘having an ethical outcome’, just procedurally obtained. But then we are also in a double-bind, as defining the outcome as ethical because the procedure is ethical is to preclude the criteria for an ethical procedure by predetermining its outcome... It seems we cannot really have a procedural, pluralistic ethics in the end.

However, and in answer to this, the complaint only works within the assumption that the outcome of ethics is itself ‘ethical’, by way of some pre-agreed upon definition of what counts as ethical. Under such conditions, emergence is unlikely and unwelcome, and generation will not occur, because these are degenerative conditions; if there is indeed only one or certain ways to be ethical, and we can tell what those are, then we already know the ends to our processes or, at least, what ‘good’ ends will be. Why then engage in the process at all and not simply present the utopian model? Above I

have argued that model-based reasoning produces, for these reasons, a different way of conceiving ethics, not just a different way of doing ethics.

Model-based reasoning as we have seen in cases such as participatory multi-modelling capture, further, a way to improve on current ethical models and tools for working with others, I argue. In such an exercise, a model can be presented as a utopia but it would play a role as a boundary object—something about which deliberation occurs—rather than as an endpoint of the process. As Nersessian (1999) points out, “it is working with this intermediate form of representation that facilitates conceptual change” (p. 15). Model-based reasoning, that is, engages the models proposed as intermediaries, as object with which to play and to manipulate, in order to gain insight and to reach a shared understanding not only of which model is best for what purpose, but also *why*. Thus, more than an agreement about principles themselves or even about which outcome is most desirable, plurality exists in why and how such worlds can and should come about. Emphasising process and using a framework such as participatory multi-modelling can open ethics practice up to transdisciplinary processes, necessary today when much of ethics is being and should be done in response to fast-moving technological, environmental and social change.

4.2 Ethics is Neither Always nor Necessarily Generative, nor Are the Outcomes of Models Necessarily Good or Even Explainable and Therefore Could Be Inappropriate for Ethics

Pointing to the situation of ethical upheaval we are contending with in our world draws out another countargument, that the goal of ethics *ought* to be consensus and working together, rather than plurality and potential polarity. I have already mentioned two issues with presuming that a good process will lead to good ends. First, creating the conditions for emerging possibilities for action and theory is not the same as creating the conditions to ensure those actions and the theories that justify them are ethical, good, or right (see also Copeland 2023). Second, without specific expert guidance or preset criteria for a good outcome, it seems there is no guarantee that even an ethical process will produce ethical results. A third related concern is whether model-based ethical reasoning will amount to a ‘black box’ process, in which the reasons why our model leads to good results remain opaque to us, contrary to our ideal of having an ethics that is not authoritarian but rather reason-based.

To begin with the third, the complexity of the world in which we do and theorize about ethics means that no matter how we model that world, it will remain incomplete. As said before, this is not a problem for model-based reasoning, but rather an intrinsic element in its application. Thus, this is not a counterargument but rather further support for turning toward process over outcome as the goal of ethics. Indeed, successful modelling as well as successful ethics demands that we return, iteratively, to our basic assumptions as well as to how we have depicted our world—as our models produce new advice on action, we thus change the very world upon which our model is based. The same should be true for ethical action, which ought to produce a more ethical world, in which future concerns will differ from current concerns. Our own history attests to this evolution, as conceptual change has led to changes in the values and ethics we hold to be good and true over time, from the turning away from our justifications for slavery to the inclusion of new identities

within our concept of humanity. And finally, together these considerations highlight that we should expect emergence and unpredictability; not only are these sources for insight and creativity, but they also characterise the impact of the actions we take in the world. That is, no matter how well we model or theorize an ethical framework, as the world changes and we change, so will the reliability of the assumptions they make. Both models and principles ought to be subject to reconsideration and change.

The second question raised is around the consequence that expert guidance is required and/or a predetermined set of criteria for success, in order to tell if a process is ethical or not. That is, the outcome needs to shape the process, and therefore comes before—a process-oriented ethics will not correct this, and so the outcome maintains its priority no matter how much we focus on the process. The simplest counter to this argument within this paper is that there is no outcome to this process by which to judge the process, and so the process must be judged by its own effectiveness, in light of its outcome when obtained. This is not to imply an only backward-looking ethics, however, as we need not spend our time evaluating past decisions for whether they are good or bad as a final judgment. Rather, the outcomes of processes become fodder for their own refinement, just as testing model effectiveness results in the refinement of the model, not setting it upon a pedestal nor throwing it away. Ethics, thus, should be seen as a continuing process without end, and not as effective or not, dependant on some outcome at some arbitrary point.

This brings us to the first question raised above: is it correct to say that ethics is generative¹, or that it should be? The implication in the phrasing offered above admittedly begs the question, but only to highlight the recursiveness of the problem itself—if ethics is ultimately about agreement, then tautologically it ought not be generative in the sense that I have promoted with this paper. But ethics is in fact generative, I would argue. Feminism, for one, has introduced the connection between epistemology and ethics, as well as developing concepts such as relationality and standpoint which have little or no presence in earlier ethical theory but are essential to understanding the contemporary world. These are not refined concepts but altogether new, having emerged from taking up new perspectives as valuable for philosophical and ethical consideration.

Further, the hidden premise in this question is that disagreement leads to conflict, and that a truly ethical world would include no conflict—otherwise, why prioritize agreement by consensus? This premise itself is an assumption, so far unproven in any interpretation of human history, I would argue. Further, and as I have already suggested, it is time to turn more toward participatory, transdisciplinary and plurality-based approaches to ethics, in the face of the actual world we live in, which is itself thus. A utopia is, recall, ultimately not a model at all but a work of art.

¹ I thank more than one keen audience member of the Keynote talk upon which this paper is based for this marvellous question, and Selen Arfini in particular for pointing me toward feminism as an illustration.

5 In Sum

This paper is not the only one to take up the correlations between model-based reasoning and ethical reasoning, nor has it gone deepest into the analytical components of that analogy. Rather, it has intended to deepen the analogy by exploring what model-based reasoning, insofar as it is a fair analogy, has to add to our understanding of not only how to do ethics but of what ethics is about and how it should be done. This normative aim suggests that further work can be done not only to draw comparisons between these two fields in which practical reasoning reflects and is reflected within theoretical principles, but to the benefit of future ethical work, in practice and in theory development. Model-based reasoning opens up avenues for shared understanding, not only among theorists but between those who theorise and those who must enact our ethical theories. That this should be a priority for philosophical ethics may be only implied by this paper, but is grounded in its appeal to new insights gained in the field of modelling, particularly where participatory multi-modelling has been employed. The aim of shared understanding for ethics, and consolidation over what possibilities there are and should be for action in the world, are practical aims for ethics but can be modelled using not only empirical observations but theoretical principles. And, given that the model itself is the subject of deliberation, these are all up for consideration, creating a boundary object but also a common goal. Plurality, that is, can exist alongside a singular shared goal, so long as we avoid conflating outcomes with finality. Model-based reasoning offers a way to do just that.

References

- Copeland, S.: Doing ethics, and the possible. *Possib. Stud. Soc.* **1**(4), 427–435 (2023)
- Cuppen, E., Nikolic, I., Kwakkel, J., Quist, J.: Participatory multi-modelling as the creation of a boundary object ecology: the case of future energy infrastructures in the Rotterdam Port Industrial Cluster. *Sustain. Sci.* **16**, 901–918 (2020)
- Edmonds, B., et al.: Different modelling purposes. *J. Artif. Soc. Soc. Simul. Artif. Soc. Soc. Simul.* **22**(3), 6 (2019)
- Epstein, J.M.: Why model? *J. Artif. Soc. Soc. Simul. Artif. Soc. Soc. Simul.* **11**(4), 12 (2008)
- Magnani, L.: Prefiguring ethical chances: the role of moral mediators. In: Ohsawa, Y., Tsumoto, S. (eds.) *Chance Discoveries in Real World Decision Making, Studies in Computational Intelligence*, vol. 30, pp. 205–229. Springer (2006)
- Nersessian, N.J.: Model-based reasoning in conceptual change. In Magnani, L., Nersessian, N.J., Thagard, P. (eds.) *Model-Based Reasoning in Scientific Discovery*. Kluwer Academic (1999)
- Nikolic, I., Warnier, M., Kwakkel, J., Chappin, E.J.L., Lukszo, Z., Brazier, F., de Vries, L., Cvetkovic, M., Palensky, P., Verbraek, A.: *Principles, Challenges and Guidelines for a Multi-model Ecology*. WhitePaper, Netherlands Enterprise Agency (1999)
- Roussos, J.: Modelling in normative ethics. *Eth. Theory Moral Pract.* **25**, 865–889 (2022)