"Ethics in the Swarm: self-coordinating opinions and emergence"

By Daniele PROVERBIO

Daniele Proverbio is the co-author of Swarm Ethics™, Director of Interdisciplinary Research at the House of Ethics and is currently affiliated to the University of Trento where he works on Complex Systems and Artificial Intelligence.

Ethics in the Swarm: self-coordinating opinions and emergence

All humans do is modelling. That is, constructing idealised and (hopefully) intelligible abstract representations of experience. Whether those representations are close to "real" and absolute truth (if any exists) or simply work well enough to keep surviving human contingent needs, is unfeasible for us to discern: we are part of the same system we try to model.

Note: this writing is also a modelling product. Hence, it builds on the *assumption* that the primary intellectual activity of humans is modelling. There may be more, but that contributes less to our scopes. Importantly, we neglect that any "revealed truth" can be immediately grasped, and keep on modelling. This circularity is unavoidable, as we speak of the same method we employ to speak.

No absolutes

From the assumption above, it immediately follows that absolute concepts are meaningless. Religions have it easy: they assume a beholder outside of the universe system, that lends the truth to a subset of chosen individuals. Shall we not agree on some of these second set of assumptions (either neglecting that such beholder exists, or neglecting that perfect communication occurs, or even that "more chosen" individuals are present), we conclude that absolutes are not accessible. That is, whatever models are widespread, they are human-made, human-spread and human-sustained; in short, social products.

This thinking applies to ethics, too. In another provocative article, "What if ethics does not exist?"[1], the idea that ethics is a model of some invented and useful set of principles is discussed and defended. Here, we further develop a model of how ethics can very well be an emerging idealisation of occurrences within coordinating social swarms, and what this entails for ethics itself and, circularly, to the swarms themselves.

Swarms

Swarms are collections of single individuals, able to perform complex tasks through self-organization. In its origin, the discipline of swarm intelligence was developed as a field of complex systems studying groups of natural or artificial entities[2]. Recently, the idea of self-organising societies[3] and emerging social models inspired a synecdoche into "Swarm ethics", to signify emerging ethics from interacting humans.

Let us clarify some preliminary concepts. "Swarms" require some simple yet key ingredients, that set them apart from mobs or masses[4]. We use a simple example, inspired by natural flocks of birds, to give some intuitive visualization.

First, swarms are composed of individuals, that interact with one another. Hence, swarms are not simply the sum of stand-alone agents, each one behaving on its own; they are more than that, since individuals are capable of sharing information, influencing each other and mutually evolve. The whole is more than the

---

[1] https://www.houseofethics.lu/2021/11/11/what-if-ethics-does-not-exist/
[2] Kennedy, James. "Swarm intelligence." Handbook of nature-inspired and innovative computing: integrating classical models with emerging technologies. Boston, MA: Springer US, 2006. 187-219.
Dorigo, Marco, et al. "Swarm intelligence." Scholarpedia 2.9 (2007): 1462.
[3] Helbing, Dirk, ed. Social self-organization: Agent-based simulations and experiments to study emergent social behavior. Springer, 2012.
[4] y Gasset, José Ortega. The revolt of the masses. Routledge, 2021.

sum of its parts. Think of flocks: they are composed by individual birds; however, they are not isolated and numb, but recognise mutual positions (very simple, yet effective information sharing). Without this ingredient, single birds would crash onto neighbours as soon as the flock divert from straight lines. This results in impressive and stunning coordination of the whole flock, while hunting or migrating.

Second, swarms are dynamic. Both individuals and their collective (the swarm) change over time, thrive and vibrate, responding to stimuli and/or internal turmoil. It is not true that, once a mass is set, it keeps on behaving the same way. Transient behaviours, changes (even abrupt) may occur. Indeed, bird flocks accept new individuals during migration, adapt to predators or sick individuals, etc.

Third, swarms are open systems. Boundaries are difficult to recognise, are fluid and ever-changing. Recognising swarms is a matter of modelling, by someone outside of it (or within, given appropriate scopes). External stimuli may influence "boundary" individuals and potentially the whole of the swarm. The environment where the swarm thrive may provide gradual or disruptive inputs. Several swarms may coalesce or compete. Atmospheric currents are accounted for by bird flocks, for instance.

Fourth, different scales can be recognised: slowly-evolving first principles, which individuals abide to and employ to set the *possibility* to connect with others and set common goals; medium-paced behaviours, entailing response to stimuli, adaptation and other actions; rapid whims or changes, fades and other transient individual modification, possibly diffusing and perturbing the swarm but quickly disappearing. Back to the flocks: "first principles" are instinctive ratios as "migrate there" or "keep distance from neighbours while flying", which set common goals and means of information sharing. Medium-paced behaviours are the mutual coordination while flocking, or the astounding acrobatics often seen. Rapid whims are, e.g., quick adjustments of birds' trajectories, like in response of atmospheric drops or of predators' hunting. Sometimes they involve just few individuals (like the air adjustments), sometimes they propagate through the whole flock (anti-predation manoeuvres) but rapidly vanish.

Fourth, circularity governs human swarms. This is key to differentiate (up to our current knowledge) societal swarms from other types or natural swarms. Through debating, social dynamics or other factors, human agents are capable of modifying – dynamically – their driving principles at all scales, on top of being driven by them. This circularity add up to the challenges of unravelling the laws of social swarms, but it is also where ethical discourse come into place.
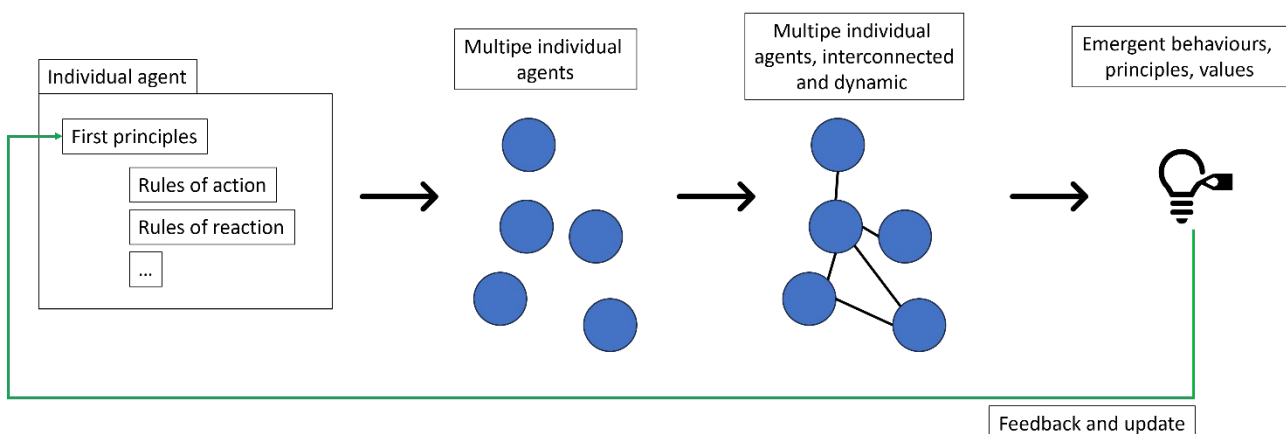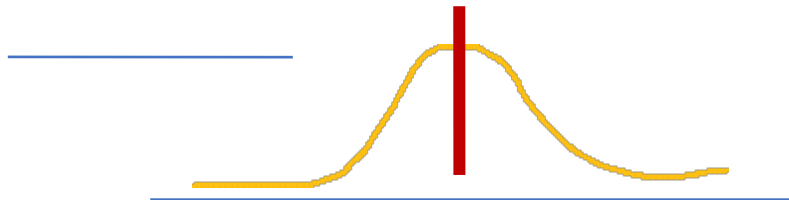

Swarm Ethics™

Building on the two premises above, we now dive into the concept of Swarm Ethics™. Here, "ethics" bears two distinct but key marks: it is recognised as a social product, an emerging model of collective principles; and it enters the swarms' dynamics in a circular way.

As the swarm thrives dynamically, beholders may recognise common patterns and measure alignment of opinions and guiding principles. They are not necessarily one and the same, but are distributed in some fashion. By studying such distribution and looking at what is more representative, the beholder may infer the existence of "mean" principles (red in Fig. below) and decide to lift them as single representative of the whole system. Here, "representative" can be statistic, vocal, marketing-wise, or other.

# of alike morals

A "distribution of morals"

Post-hoc, one can call these mean guiding principles "ethics", granting special status and electing them as general beacons. By feeding back this information into the swarm, through means of communication, such "ethical" principles add up to the distribution of principles, eventually skewing it like "self-fulfilling prophecies". Once becoming part of the set of underlying principles and values, ethics drive swarms by providing interaction purposes, shaping goals and constraints, and overall modifying the slowly-evolving basics.

This process could subtly proceed as long as few "beholders", or thinkers or opinion-setters shaped Western culture, and as long as information sharing was slow. The modern age brings radically different perspectives. First, it opened the breadth of potential swarming participants, by connecting the world. Swarming networks become bigger, more heterogeneous and dynamic. Second, it drastically sped up information sharing, thus disrupting the time-scale separation between basic values, behaviours, fades and trends. They are all at risk of mixing up. Third, it brought an ever-changing intersection of values ad principles. More than relativism (where multiple ethical sets are recognised but remain separate, like oil and water), the modern era is characterised by mixing and fuzzy boundaries (like pouring sugar into water and never separating them again). Globalization, world-wide travels and business, the Internet – it all concurred[5].

---

[5] Barabási and Bonabeau. "Scale-free networks." Scientific american 288.5 (2003): 60-69.
Percacci and Vespignani. "Scale-free behavior of the Internet global performance." The European Physical Journal 32 (2003): 411-414.

Historically:

Single thinkers (hubs), behaving as "trend setters" and "ethical models" (e.g. the Curch, Kant…)

or

Few links and small networks; slow spreading of information

Digital age:

Giant world-wide networks
- Decentralised
- Hubs exist, but most networks are scale-free (e.g. the Web)
- High speed of info spreading
- Dynamic pluralism (rather than relativism) of thoughts and ethic-es

Some consequences

Listing and discussing all consequences of Swarm Ethics™ is not only unfeasible, but also undesirable, as new ideas will necessarily come dynamically from feedback and discussion among other researchers and practitioners. Other relevant consequences and influences to philosophical constructs, translational applications and business alike can be found in the various resources collected on the Swarm Ethics™ website[6]. We here consider three macro-consequences, pertaining to different domains.

1. No first-principled ethics, but post-hoc recognition, mutual discussion, feedback into the system. At first, swarm ethics look descriptive instead of prescriptive, since it begins by breaking down consequential and traditional ethics and provides a new social and anthropological description of ethics. However, it is not "just" descriptive. The inherent circularity of the process, to which Swarm Ethics™ belongs as modelling attempt, feeds it back into the system. The post-hoc recognition that first principles are non-necessary opens a range of possibilities to the definition of values through mutual discussion and information sharing propelled and incentivised by swarm ethics ideas.

2. One driving principle: inclusion into swarms, equalities of human agents, openness to dialogue. From point 1, it follows that the only prescription, which is a direct consequence rather than an absolute imposition, is: let the swarms thrive. Isolation, concentration of information power within few nodes, discrimination – they all stagnate and constraint societies. All agents should have equal possibilities and the potential to contribute, so that evolution may proceed and bring forth novelty.

3. Embracing modern and complex challenges at the price of circular arguments: need to continuously envision and develop analysis and intervention frameworks. The "easy way" of setting absolutes crumbles in recognition of dynamical emergence. Simple feed-forward ruled-based interventions are less effective. Instead, the capability of involving into dialogues, and to invent and generate slowly-varying basic and shared principles, beyond relativism, could drive behaviours and fast reactions. Circularity makes this process more arduous, but at the same time more robust.

Conclusion

A complex world requires complex ideas. The linear thinking inherited by Enlightenment and Ancient philosophy suffers in face of dynamical, open and uncertain human exchanges. Swarm Ethics™ paves the

---

[6] https://swarmethics.com/

way to new descriptive and guiding frameworks, capable of fostering the emergence of novel and modern ideas to tackle the modern challenges.