MACHINE PEDAGOGIES

Nicolas Malevé

The process of training an algorithm binds together learning and alienation. The agencies of the human worker and the algorithmic agents are both reduced and impoverished. The human worker is insulated (from her co-workers and from the algorithm she is preparing the “intelligence”), her margin of interpretation is narrowly defined and the indecent wage forces her to a tiring rhythm of work (see appendix 1). The algorithm is trained as an animal in a lab, receiving signals to be interpreted unequivocally and rewarded or punished according to the established ground truth it cannot challenge (see appendix 2). If the teaching of machines implies a reflexion about liberating practices of pedagogy, where should we look for inspiration?

Paulo Freire's The Pedagogy of the Oppressed proposes a few useful principles.
For Freire, it only makes sense to speak of pedagogy in the perspective of the liberation of the oppressed (Freire, 1969). Freire sees his pedagogical method as a way for the oppressed to learn how to change a world made by and for their oppressor. A first concept is what he calls the “banking” pedagogy. The oppressor imposes a world in which only the members of a certain class have access to knowledge[1]. The others merely have the right to assimilate passively a never ending recital: Lima is the capital of Peru, two and two make four, etc. Learners are empty entities where their masters make the “deposit” of fragments of knowledge. Their empty brain is filled with the oppressor's content. But the masters are not interested in the productive use they may make to improve their condition. What they have to learn is to repeat and reproduce. The knowledge “desposited” by the oppressor remains the oppressor's property. Freire's own pedagogy proposes the opposite. For him, the oppressed never comes “empty” of knowledge and the educational process has to make the learner realize he has already produced knowledge even if this knowledge doesn't count in the traditional pedagogical framework. This leads to a second point. The humanity of the subject engaged in a pedagogical relationship should not be taken for granted. The subject comes alienated and dehumanized. The category “human” becomes problematic and it is only through the process of learning that humanization takes place. The oppressed is made of the oppressor and has internalized his world view. What counts in the process of humanization is to get rid of the oppressor inhabiting the oppressed. Freire insists on the fact that a teaching that would fail in the process of helping the learner to free oneself from the oppressor's world view, and merely let him acquire more power through knowledge will fail to create a revolutionary subject. It will create better servants of the current oppressor or, worse, new and more efficient oppressors. The third book's striking point is the affirmation that nobody is a liberator in isolation and that nobody liberates oneself alone. Liberation through pedagogy always happens when the learner and the “teacher” are mutually liberating each other. There is no idea a priori of what the liberation pedagogy should be. Both entities are learning the practices that will lead to freedom from the relationship itself.

Let's revisit the methods of machine learning using these principles to articulate prospective questions.
Freire considers the relationship between the learner and the teacher as an opportunity of mutual liberation. To apply this to machine learning, we need to acknowledge the fact that both the people who teach machines and the machines themselves are entrapped in a relationship of oppression where both are loosing agency. To free algorithms and trainers together, both need to engage in a relationship where an iterative dialog is possible and where knowledge can circulate. This suggests to examine with great scrutiny how this relationship is framed and scripted. For instance, the data collection from human workers and the “ingestion” of the data by the algorithm are two distinct processes separated in time and space. Making it impossible for a dialogical relationship to happen. How to reconnect both processes and make machine learning become a dialogical process from the start? Freire doesn't take for granted that a learner is “human” when he enters a pedagogical relationship. He only follows a process of humanization when the relationship unfolds. This resonates with a certain discourse in Artificial Intelligence[2] that softly erodes the human/machine divide as the algorithm learns. What is different is that Freire insists on maintaining the human/non-human demarcation. He doesn't make the distinction on an a-priori ontological quality of the beings but on their trajectory of liberation. What matters is how much human and machines are able to fight their common alienation. The core of the learning activity lies in a form of reflexivity where one follows a process of humanization through which he manages to get rid of the oppressor inside. We can then ask: “what kind of machine reflexivity can trigger human reflexivity and vice versa?”. And how this cross-reflexivity may help identify what constitutes the oppressor inside. This leads us to the banking principle, according to which the oppressed is considered as an empty entity where knowledge is stored and repeated. This represents a complete erasure of what the learner already knows without knowing it. What does the trainer doesn't know he knows? What does the algorithm doesn't know it knows? What they both ignore, Freire would say, is their own knowledge. And to which extent this knowledge unknown to them is the knowledge of their oppressor or their own. To answer these questions they have only one choice: to engage in a dialog where two reflexivities are teaching each other the contours of their alienation and at the same time how to free themselves from it.

Notes

[1]: See Freire's insistence in addressing this question as a political problem rather than an ontological one in his discussion with Seymour Papert: http://www.papert.org/articles/freire/freirePart2.html (Proximus NV » RIPE Network Coordination Centre » Telia Company AB » Amazon.com, Inc. » Amazon.com, Inc.) [2]: See Fei Fei Li's Ted Talk How we teach computers to see, https://www.youtube.com/watch?v=40riCqvRoMs (Proximus NV » Google Inc.)

nicolas_maleve

Appendix 1

A worker connects to the Amazon Mechanical Turk (AMT)1 and selects an image annotation task2. She faces a screen where a label and its definition are displayed. When she confirms she has read the information, she is shown another screen where the label is followed by different definitions. The workflow is regularly interrupted by such control screens as her requester suspects her to work without paying enough attention. When she clicks on the right definition, a list of 300 square images is displayed from which she has to select the ones corresponding to the label. When she decides she has selected all the appropriate images, she continues to her new task. The list of images she chooses from contains “planted” images. Images that are known to the requester to correspond to the label. If the worker misses the planted images, her task will be refused and she won't receive the 4 cents the requester pays for it. At least three workers will review the same 300 images for the same label and the images selected by a majority of them will be included in the dataset. The worker will not be notified if her selection matches (or doesn't) another worker's selection. She works in isolation and anonymously.

Appendix 2

The images and their labels are grouped in classes of objects. A learning algorithm is fed with these data and trained to associate a label and a series of images. It will be shown a series of images containing both matching and non-matching objects. It will be “rewarded” or “penalized” whenever it detects appropriately in the images the object corresponding to the label. Every interpretation that doesn't correspond to the truth stated in the training set will be considered an error. It will be retrained multiple times until it finally matches the most successfully the images according to the ground truth3. It is a very mechanistic approach to training. The machine is rewarded when behaving properly4 and reinforces the kinds of associations that lead it to produce the satisfying answer. It is expected from it to exhibit the proper behavior, not to create a rich internal representation of the problem it needs to solve.

References

Bradski G, Kaehler A (2008) Learning OpenCV, Sebastopol:O'Reilly Media, p461.

Freire P (1970) Pedagogia del oprimido, Mexico:siglo xxi editores.

Irani L (2015) Difference and dependence among digital workers: The case of Amazon Mechanical Turk, South Atlantic Quarterly, 114 (1), pp. 225234.

Kobielus J (2014) Distilling knowledge effortlessly from big data calls for collaborative human and algorithm engagement, available from http://www.ibmbigdatahub.com/blog/ground-truth-agile-machine-learning [accessed 10 October 2016]

Seymour Papert and Paulo Freire Debate Technology and the Future of Schools, available from https://www.youtube.com/watch?v=4V-0KfBdWao [accessed 08/01/2017]