Benjamin Eckhardt (, Alisa Andos, Arnel Pällo)

For finding somehow novel and usefull, aka creative, solutions with evolutionary algorithms one of the most important parts is the formulation of the solution space. A well-defined solution space can prohibit the exclusion of viable solutions and can enable focused exploration only in meaningful dimensions.

When we create, we think, we all the time form hypothesis about what kind of structures and relations could be meaningful ways in which to view the solutionspace. We assume we all the time have multiple partial and overlapping dimensions to represent the things in our head, quickly proposing new ones, and switching a lot. We then do random alternation, but only in those meaningful dimensions. The hypothesis are highly informed by previous hypothesis - formed analytically out of them, so that we don't need a lot of experiments to validate. And the hypothesis are not just on one level of structure (like 'oh this melody sounds good') but can regard any possible kind of relation: any complex relations between structures (melodies are constrained in certain ways by the harmonies that they cling with) and relations between relations (when melody and harmony in one bar behave like this, in the other the most promising relations to explore are that).

So we would want to build a thing that: Creatively proposes hypothesis for meaningful relations on any level, building on each other, and then creates by alternation of the parameters (to find new hypothesis for new relations). The relations are the space in which to propose solutions, but there need to be multiple spaces all the time. As to come up with those new framings informed by experiences in the previous framings may be the actual core of (some kind of) creativity.

As a possible approach to this problem of automatic solution space formulation, we consider to research possible ways of coevolving an archive of multiple solution spaces, which are supposed to be used by a coupled second EA to drive exploration by combining and alternating between them. (This is maybe touched by [arxiv.org/abs/2105.00682, arxiv.org/abs/2105.10317 ?].) The embedings could be described by autoencoders, though they naturally need lots of samples, which we object, as our creativity does not need them. Though they might turn out unsuited for dynamically evolving.

The process could leverage a multi-level archive, by which we hope to capture the above recursive relations-between-relations-hypothesis with (though we have to think further on it). Maybe we find a way to express this recursion via a growing tree-like structure rather then a fixed height stack of archives. This in it alone appears unprecedented to us? We could thereby test the hyptothesis, if the effect of performing exploration in varring combinations of a quite high number of very low dimensional feature spaces, rather than one very high dimensional with lots of meaningless recombination possibilits, is comparable with the computational benefits of convolutional or transformer architectures vs fully connected networks?

If that turns out fruitless or not enough, we can plugin mechanisms like human-colaboration [arxiv.org/abs/2003.03377, arxiv.org/abs/2003.03268] or gradients and backpropagation [dqd-rl.github.io/].

One core question definitely seems to be: How to characterize 'meaningfull' ways of representation when we don't wont to hardcode too much. We shall find something domain generic, maybe statistical properties of solution spaces, or learning from the primary source of 'meaningfull' that we wanna adress anyway: human cocreators [like in arxiv.org/abs/2003.03377, arxiv.org/abs/2003.03268]