Q&A
10 November 2021
Quantum Machine Learning
Professor Anatole von Lilenfeld (University of Vienna)

Q1: With the ability to go from one element to another and do the alchemical transitions. Presumably you grow atoms as well by just going from a zero-charge atom to some nuclear charge on it, to grow in atoms to make your molecules bigger. Is that something that the set will work on?

We actually did this, I think we published it two years ago, you take a box with a uniform electron gas and the background charge, and then you reduce that background charge, and you can grow individual atoms. And then if you apply the chain rule you can even print out the atomic contributions to the total energy, so you get a measure of the Atomic Energy, and you can even also with the same trick, get the contribution to the total electron density. So, you also get atomic electron densities in your molecule.

So, if you take your box and make it a shape, that might happen to be the active area of an enzyme or a drug or something, of a drug target and have an external field there that's represented by the thing you're trying to bind to, can I grow a potential binding drug?

I don't see why not? My main concern would be that you would be tied to this particular configuration of your external field and so on. Maybe some averaging might be helpful or to do this one ensemble of configuration. But in principle, yes absolutely this should be possible. I don't see any physical reason why not.

Yes because, the feedback between what you're growing in and the protein that would have to come from the ensemble.

Q2: In the fragment system, you have these fragments and you've shown how representative of the whole of very reasonable accessible chemical space you can reproduce really quite accurately. With the generative models, can we turn that around and see if we could generate molecules which are not simply obtainable from those fragments? So, find that new novel, the C60's of the world etc., suggestions to work in places of chemical space that are not in fact representations. In a sense, finding new fragments, is what we're talking about.

*Yes, that's an interesting Point. I won't exclude it, but I think the chances are minimal. If you think about it in terms of an alpha principle or so, right where you go, you start with an atom and you try to exhaustively list all it's possible chemical environments. Then at some point, you've exhausted it and that's it and then you build the next shell, and you do that exhaustively, etc. So of course, at some point the system will be so large it also starts to explode, combinatorial, and then maybe there might be this into this region where it's so large you cannot do it exhaustiv*ely*, so then you could use a generative model to possibly find the most relevant fragments. I think that that would be the most likely scenario where this could work, but for these really small fragments I think it's quite finite actually.*

I think the evidence you present is that with the number of very reasonable number of fragments you are able to cover a lot. What we have experimentally accessed so far you seem to predict very well.

Mind you, it's just CHON (Carbon, Hydrogen, Oxygen and Nitrogen) so just the few elements from organic chemistry. So, as we start adding the later elements in the main groups or some transition methods and so on, the space will we still grow. but it it's still finite, I think it's very feasible.

Yeah, yes. I mean you're right. We do have another sort of 100 odd elements that we could in theory use, and then the relativistic corrections will start to change some things and then we need new stuff. But there's not so many experimental data to put in to do the calculation, so that becomes more of a problem. At that point.

I think this this might be really some fantastic challenge for humanity to sort of have a systematic experimental data set on those little fragments from which we could then build projected or extrapolating models.

Well, I think that's something where we are looking at this the other way around. Asking the question how many of those fragments do we not have reliable bits of information on? I think that would be a really interesting list to publish for the synthetic chemists to target making interesting molecules from those fragments to provide exactly that ground truth. You always worry if you don't have some ground truth somewhere, there's always a worry that something's gone wrong.

I find it shocking how scarse the literature looks like in terms of data, so I once asked a postdoc to collect melting points for inorganic systems, and he could only find less than 2000 examples in the literature of different systems, so that's what we have.

It's not enough, considering how many elements that is.

it's very sobering.

Thank you. And I must agree with you that the one of the exciting things about this community is the fact that the data and the algorithms are shared so that you know you can work out whether you've understood it because you can run it and see what going on and move on from there and that is something that is particularly marked in this community, which I think is another reason why there's been so much progress.