

Computational Molecular Biology 2019

Final Assignment: A Variation on AlphaFold.

Due Tuesday June 4th 2019, 23.59h

Write down your answers for this final assignment in a *.pdf* file with the following name “<your student number><your last name>_final.pdf”, e.g., “012345jansen_final.pdf” and send it as an attachment of an e-mail with subject “CMB2019_final” to erwin@liacs.nl.

Note: Groups of 1 – 2 students are allowed. In case of a group of 2 students, be sure to mention all student numbers, names and affiliations.

Description

De novo structure prediction is one of the most challenging tasks among the Biennial CASP Experiments. Starting with CASP11 we saw the emerging successes of so-called co-evolutional methods, where evolutionary coupled residues detected by multiple sequence alignment of protein sequences give valuable information for further constraints, distances etc. The co-evolution of these residues suggests physical couplings of some kind, may it be proximity or others. Protein structure prediction methods predicting and using binary contacts and inter-residue distances [2, 3, 5] showed impressive performance improvements during CASP11, CASP12 and CASP13.

The winner of the CASP13 [1] de novo structure prediction experiment is AlphaFold by DeepMind [2, 3] it uses neural networks for inter-residue distance predictions and angle predictions based on evolutionary coupled residues found by MSA (see Figure 1.).

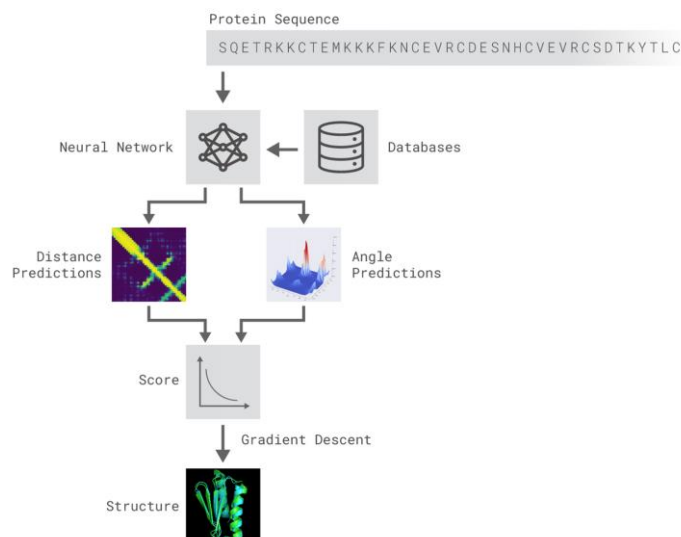


Figure 1: Overview of AlphaFold by DeepMind from [2, 3].

In this assignment you are asked to propose, describe and substantiate a variation on the AlphaFold algorithm that would be interesting to research further as a contender for CASP14. You should describe your proposal in a 4 page document, structured as a technical report: abstract, introduction, relevant state-of-the-art, your proposed variant, argumentation and substantiation of your proposal, proposed implementation, experimental setup to validate your proposal, conclusion and discussion, and the list of references.

Note, it should be a proposal only, you do not have to implement or experimentally verify the validity of your proposal.

Reference [4], on which many slides of the last lecture are based, is a very inspirational piece of writing on AlphaFold, e.g. ‘The way forward’ [4].

References

- [1] CASP13: <http://predictioncenter.org/casp13/index.cgi>
- [2] R.Evans, J.Jumper, J.Kirkpatrick, L.Sifre, T.F.G.Green, C.Qin, A.Zidek, A.Nelson, A.Bridgland, H.Penedones, S.Petersen, K.Simonyan, S.Crossan, D.T.Jones, D.Silver, K.Kavukcuoglu, D.Hassabis, A.W.Senior, “De novo structure prediction with deep-learning based scoring.” In Thirteenth Critical Assessment of Techniques for Protein Structure Prediction (Abstracts) 1-4 December 2018. Retrieved from [here](#).
- [3] <https://deepmind.com/blog/alphafold/>
- [4] <https://moalquraishi.wordpress.com/2018/12/09/alphafold-casp13-what-just-happened/>
- [5] Jinbo Xu, Distance-based Protein Folding Powered by Deep Learning, Nov. 2018. (<https://arxiv.org/abs/1811.03481>)