Chemical and Bio Synthesis use-cases: Edible Polymers, Carbon Capturing, and CBD Isolate Phytocannabinoids

☑ **Author's Note:** In regards to a computationally derived pipeline for synthetic processes.

Andrew M.K Nassief
Genomics Researcher at Bleunomics
Quantum Engineer at the Lonero Foundation
58757 #100 VAN DYKE RD
WASHINGTON, MI 48094-9998

February 18, 2021

Abstract

Computationally artificially derived chemical synthesis and multistep preparation processes can have a variety of use-cases. The idea of utilizing advanced computational complexity, bleu-score computational ranking type systems, and isolating different variable data is integrated in the synthesis preparation processes. Some use-cases presented include: edible polymers and the field study of printable foods, synthetic carbon capturing polymers and biomasses, and CBD Isolate for phytocannabinoids. Synthetic chemistry can tackle a variety of issues from food shortage to the opioid crisis. A computational scoring system can help with elimination reactions and non or organic reactionary mechanisms. Also, accuracy through validation-scoring and forms of molecular mechanics is crucial in the experimentation process. The same applies to various forms of variances, scattering and composition accuracy. The easiest way to demonstrate some of these concepts is through the analysis of said systems for these usecase varieties.

Keywords: *Chemical Synthesis, Biochemistry, Biosynthesis, Computational Complexity, Bleu, Bleu-Scoring, Cross Validation, Sample Testing, Phytocannabinoids, Carbon, Carbon Neutrality, Organic Chemistry, CBD Isolate, Polymers, Polymer Chains, Molecular Mechanics, Composition Accuracy*

Research Purpose

The purpose of this research is for demonstrating the algorithmic and mathematical necessities in regards to the proposed chemical synthesis simulation and preparation process. The secondary purpose is to demonstrate through said use-cases. The third purpose is to demonstrate the importance of said use-cases.

Conflict of Interest: *Currently, there is no conflict of interest. This research was primarily self-funded and not done through a grant*.

Introduction

Three target areas of synthesized chemical reactions for preparations are: edible polymers, synthetic carbon capturing polymers and biomasses, as well as CBD isolate for phytocannabinoids. These are what will be focused on in regards to this research, and are considered "hardness" problems in regards to synthesis and sample complexity. Likely the best way to tackle said complexities is through some sort of computational derived method. Cross-validation, ranking/scoring, and isolating variable data computationally for synthesis preparation are likely the key areas needed to be tackled. Folding crossing mechanisms and likely some sort of distributed computing pipeline can help with such a method. It is also important to look at the raw data through some sort of validation scheme that allows the scoring system a high degree of accuracy. This is related to all forms of variances, composition, and categorized matching schemes.

Target Areas	**Complexity**	**Ranking**	**Accuracy**	**Matching**
Edible Polymers	Length	Bleu/Scores	Cross Validation	Variable Data
"CC" Polymers & Biomasses	Synthesis	Ease of Preparation	Fold Crossing	Manifolds, MMs, and Processes
CBD Isolates for Phyto...	Molecular Composition	Predicted Hardness	Variances	Category Schemes

Table #1: Pipelines and Experimental Processes

Synthetic Chemistry:
Computational Pipeline 1.0

Algorithmic Structure

The algorithmic structure most suitable for the computational pipeline will likely be broken into two parts. Those two parts are: Complexity and Processing. Complexity is also broken up into two core categories. The chemical complexity, and the computational complexity based on the chemical complexity. You might run different simulations in regards to reactionary dynamics and potential response variables in order to create an optimization predictive pattern. Given potentially heavy computational processing, one needs to figure out the most efficient way to process said hardness problems (complexities).

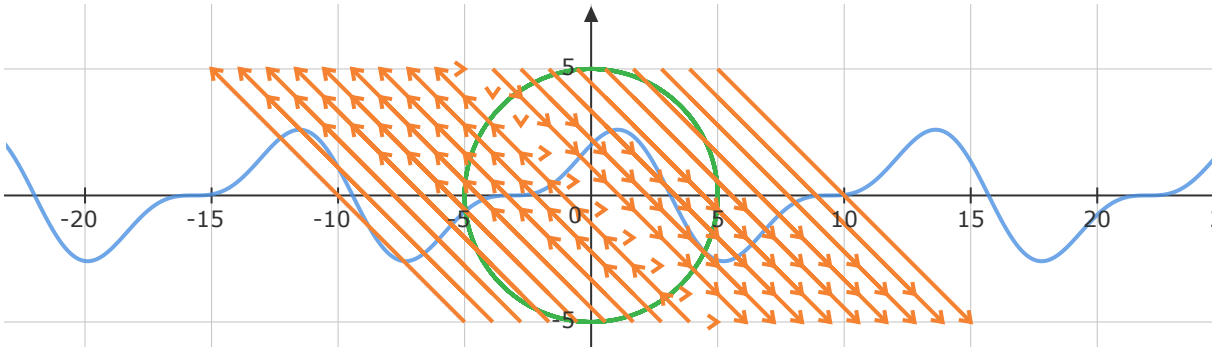

Figure #1: Time Series Optimization Plot (Wave Illustration)

The above is for illustrative purposes to demonstrate wave variance targeting mechanisms for predictive optimization models. The higher the degree of accuracy in regards to processing the complexity, the smaller the center plot for "potential variables". Below is a UML diagram of the potential processing pipeline:

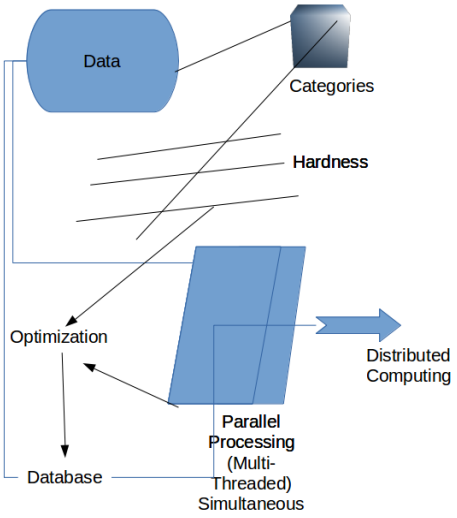

Figure #2: Algorithmic Pipeline Diagram

First you start with the data which is likely raw molecular data, diagrams, structures, variants, etc. An example of data sources (for demonstrative purposes) is ChemSpider: http://www.chemspider.com/ or chemical data on Figshare's Chemistry category: https://figshare.com/categories/Chemistry/38. The next part of the pipeline is the categories in regards to complexity. While you have the hardness optimized, processing is done in multi-threaded simultaneous processing running on parallel and distributed computing architecture. A database integrates the results, and the algorithm reruns as needed until you have the results fully optimized in the database.

Mathematical Logic

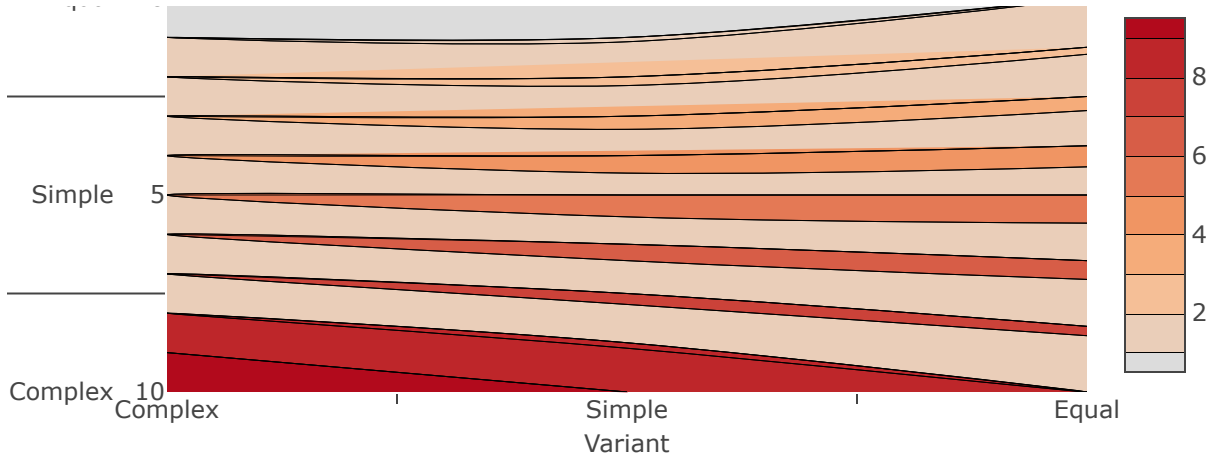

Figure #3: Complexity Scale (Heatmap) for Illustrative Purposes

Equation #1

⇒≡¬∃! // If equal, then don't continue as this isn't unique. There only exists one.
⇔∃𝔻∨¬⇒≡¬∃! // If and only if there exists a domain of values, or not the previous equation
∀⇔∃𝔻∨¬⇒≡¬∃! ⊤∨F:⇔ // whether values are matched or unmatched, please define
The further the steps, the higher the complexity.

Something to emphasize is that the higher the complexity is, the further the steps. The first step in regards to looking at proposed datasets and proposed outcomes, is whether they are equal. If they are equal, why would one run a pipeline? If not, then you need to find the proposed values in regards to the reactionary mechanisms. Once one is failing in regards to matched accuracy, predictive optimization is in place until there is a match.

Ranking and Accuracy

In order to come up with valid outcomes in regards to better optimization, ranking and accuracy are very important**[1]**. For ranking, we have the categories: Bleu/scores, ease of preparation, and predicted hardness. For accuracy, we have the categories: Cross validation,

fold crossing, and variances. BLEU stands for Bilingual Evaluation Understudy[2], and scores machine learning translated texts from one algo to another. BLEU scores can also be used as a standard metric for calculating text translations against human translations[3]. A way to implement it in regards to the computational pipeline, is scoring logic translation metrics, and category translations for human designed small molecules as well as sentiment. Likely, this would need to be implemented with some sort of precision metric in order to increase categorical ranking efficiency. The secondary category in regards to the ranking pipeline section, ease of preparation, is also quite critical. This categorizes various forms of stress tests for output or n result, while predicted hardness is on the basis of past computational processing actions.

In regards to accuracy, cross validation is also important. Cross validation in this setting, helps evaluate ranking estimator performances[4]. It is also important to put an emphasis on fold crossing. You are creating subsets and knowing how to categorize sections of the data[5]. Next is variances, which is a crucial part of the accuracy section of the pipeline. Variances allow you to look at discrepancies from the sample size and mean value [6]. You are simply creating various markers for your data in regards to the accuracy section/model.

Matching

Matching is also a crucial component in regards to a proposed computational synthetic chemistry pipeline. First one wants to look at the resulted variable data. What are the core variables? What can or has been ranked or scored? What can be pushed further? What variables match, and what are potential reoccurring outliers? One also wants to look at chemical manifolds or general applicability of said rate constraints in regards to the variables[7]. Things such as modeled molecular mechanics and simulated chemical processes are also crucial. One also wants to consider categorical matching schemes.

Target Areas (Data) → Complexity → Ranking and Accuracy → Matching

An algorithmic pipeline should have all sections working together. While there will likely be a certain order of things code and structure-wise, this isn't necessarily telling the full story. This is in part due to the fact that for reoccurring steps, predictive modeling may optimize some steps they may need to be performed at a different order or some steps needed to be done and some that don't. This is why establishing the mathematics and logic behind the predictive nature of the pipeline is fairly important.

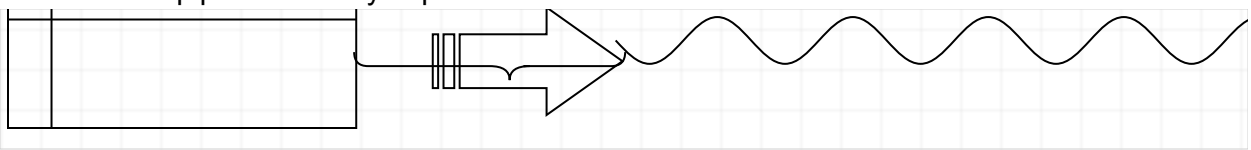

Figure #4: Data Pipeline and Waves Illustration

The above figure is primarily for demonstrative purposes. It is meant to illustrate data coming in, computational processing and the algorithm running. It is partially in reference to *Figure #2.*

Synthetic Chemistry:
Computational Pipeline 2.0

Target Areas

Now with edible polymers, there may be forms of organic synthesis processes or artificial non-organic processes needing to be derived in the reactionary process. Chemical reagents or reactants can be both a causation and testing catalyst**[8]**. For example, acetic ahyndride is an organic synthesis regent. The reactionary process is especially critical in edible polymers and plastics. 3D printing food is likely to become a thing in the future**[9]**, with unknown flavor variants that would need to be chemically compatible with the edible polymers.

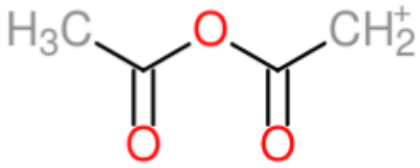

Figure #5: Acetic anhydride (Organic Synthesis) Reagent

Now with edilble polymers, there may be forms of organic synthesis processes or afticial non-organic processes needed to be derived in the reactionary process. Chemical reagants or reactants can be both a causation and testing catalyst**[8]**. Acetic ahyndride is an organic sythesis regent for example. The reactionary process is especially critical in edible polymers and plastics. 3D printing food is likely to become a thing in the future**[9]**, with unkown flavor variants that would need to be chemically compatible with the edible polymers.

Carbon-Capturing Polymers		
Binding	Molecular Structure	Molecular Dynamics
Activation	Reactionary Process	Molecular Mechanics

Table #2: Carbon Capturing Polymer Considerations

The other target area that I want to look at is carbon capturing polymers, which is important in "reversing climate change" or the common goal of carbon neutrality. Carbon capturing polymers have various CO_2 confining qualities or some sort of coordinated reactionary variable**[10]**. In fact table #2, takes into account important considerations in regards to the carbon-capturing polymer discovery process, as well as its properties. Other areas of research can include synthetic biomass and bio-energy products**[11]**. Even synthetic polypeptide reactions within algae**[12]** or coral reef net sinks**[13]** could be potentially viable solutions.

☑ **Author's Note:** This is a field I am experimenting in through proprietary means, and have had somewhat promising results.

An important consideration in regards to carbon capturing polymers, are the same type of complexities the proposed computational algorithmic pipelines would be targeting.

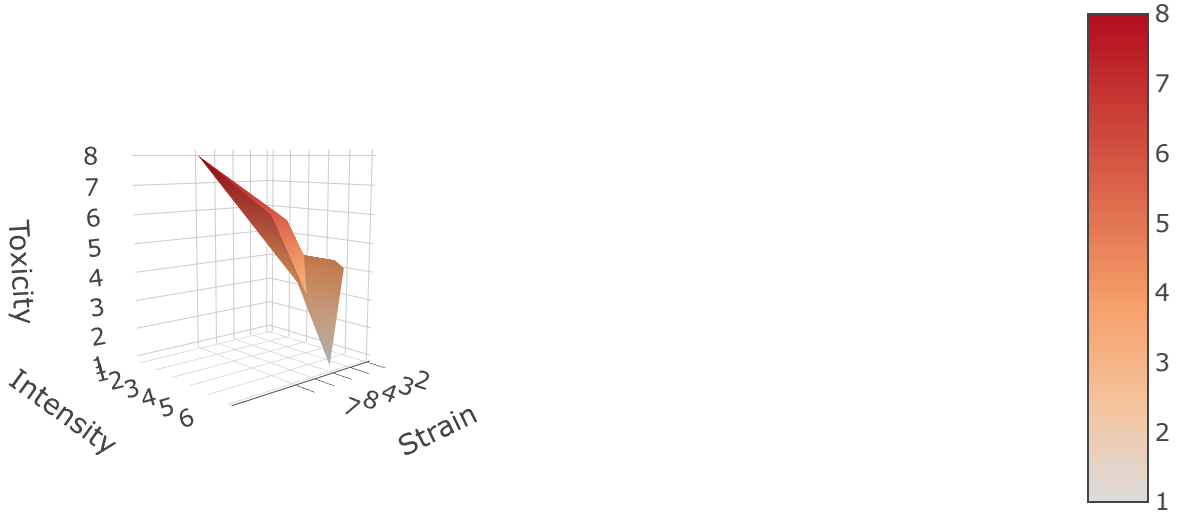

Figure #6: CBD Strain Toxicity Considerations (Illustration)

The third area I want to target in regards to solvable or investigatable use-cases, are CBD isolates for phytocannabinoids or synthetic CBD strains. Two core reasons this is important is because the opioid crisis, as well as the need for safer medical marijuana or CBD alternatives. CBD can have somewhat harmful side effects, especially in children prior to full scale brain development, or side effects such as nausea or fatigue**[14]**.

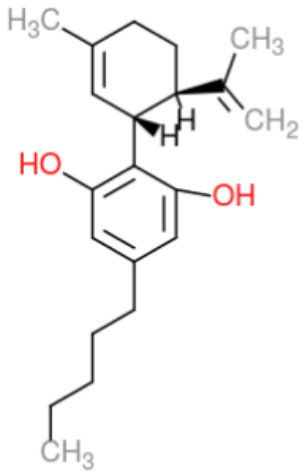

Figure #7: Cannabidiol ($C_{21}H_{30}O_2$) Structure

The above references the structure of Cannabidiol using ChemDraw and the PubChem CID**[15]**. Cannabidiol is an example compound of a drug with still highly toxicable properties. This is true even considering lesser psychoactivity and non-hallucinogenic properties in comparison to THC**[16]**. Also many biological effects and/or molecular mechanisms, or human response variants, are yet to be fully known.

Proposed experimentation can be centered around removing toxicity or identify regions of toxicity within strains. Separating the healing properties from the toxic is quite crucial.

Synthetic Chemistry:
Computational Pipeline 3.0

Process

A good example of a dynamic modeling software that can be used for chemistry pipelines as well is PottersWheel**[17]**. The PottersWheel software does have somewhat CPU-intensive functions and operates various forms of refinement and prediction processes**[18]**. There are various other forms of chemical-process modeling**[19]** and this is outside of the algorithmic pipeline and runtime. The computational/algorithmic pipeline processes by offloading data through multi-threaded simultaneous parallel processing and a shared grid computing system. The system will have logical constraints, but be self-optimizing within those constraints.

Optimal Scenarios

The most optimal scenario is getting a large variety of quality matches for the reactionary process, and modeling the synthesis in the quickest way possible with highest degree of accuracy. Accuracy in this instance, means certainty in regards to the accuracy considerations and overall computational pipeline.

Conclusion

Building a computational pipeline can help you find different complexities in regards to the chemical synthesis process for a variety of targets. The algorithmic structure of the pipeline utilizes mathematical logic specifically built for that purpose as a whole. As more and more data is extracted, the higher the degree of statistical accuracy in regards to the entire process. Three use-cases which include: Edible polymers, carbon capturing polymers and biomasses, as well as CBD Isolates for phytocannabinoids, have been mentioned in regards to target areas that can benefit from such a pipeline. Experimentation or algorithms are currently being doing in all three target areas. Accuracy is important because it helps create plentiful and beneficial experimentation and data. Computationally heavy tasks are integrated to a custom architectural framework centered around simultaneous multi-threaded parallel processing and grid computing networks. Optimization is also integrated to the computational processing and hardware architectural decisions. This helps offset CPU-intensive and QPU tasks.

References

[1] anthr. Metrics for evaluating ranking algorithms. Cross Validated (2015). Available at: https://stats.stackexchange.com/questions/159657/metrics-for-evaluating-ranking-algorithms. (Accessed: 17th February 2021)

[2] Brownlee, J. A Gentle Introduction to Calculating the BLEU Score for Text in Python. Machine Learning Mastery (2019). Available at: https://machinelearningmastery.com/calculate-bleu-score-for-text-python/. (Accessed: 17th February 2021)

[3] Mao, L. Bilingual Evaluation Understudy (BLEU). Lei Mao's Log Book (2019). Available at: https://leimao.github.io/blog/BLEU-Score/. (Accessed: 17th February 2021)

[4] 3.1. Cross-validation: evaluating estimator performance. scikit Available at: https://scikit-learn.org/stable/modules/cross_validation.html. (Accessed: 17th February 2021)

[**5]** Schneider, J. Cross Validation (1997). Available at: https://www.cs.cmu.edu/~schneide/tut5/node42.html. (Accessed: 17th February 2021)

[6] Definition of variance - Chemistry Dictionary. Chemicool Dictionary Available at: https://www.chemicool.com/definition/variance.html. (Accessed: 17th February 2021)

[7] Matthias Feldmaier, Philippe Schraft, Robin Bardakcioglu, Johannes Reiff, Melissa Lober, Martin Tschöpe, Andrej Junginger, Jörg Main, Thomas Bartsch, and Rigoberto Hernandez The Journal of Physical Chemistry B 2019 123 (9), 2070-2086
DOI: 10.1021/acs.jpcb.8b10541

[8] Reagent. Wikipedia (2020). Available at: https://en.wikipedia.org/wiki/Reagent. (Accessed:
17th February 2021)

[9] Carolo, L. 3D Printed Food: All You Need to Know in 2021. All3DP (2021). Available at: https://all3dp.com/2/3d-printed-food-3d-printing-food/. (Accessed: 17th February 2021)

[10] Wu, P. et al. Carbon dioxide capture and efficient fixation in a dynamic porous coordination polymer. Nature News (2019). Available at: https://www.nature.com/articles/s41467-019-12414-z. (Accessed: 17th February 2021)

[11] Bio-energy with carbon capture and storage. Wikipedia (2021). Available at: https://en.wikipedia.org/wiki/Bio-energy_with_carbon_capture_and_storage. (Accessed: 17th February 2021)

[12] Lamm, B. Algae might be a secret weapon to combatting climate change. Quartz (2019). Available at: https://qz.com/1718988/algae-might-be-a-secret-weapon-to-combatting-climate-change/. (Accessed: 18th February 2021)

[13] Kinsey, D. W. & Hopley, D. The significance of coral reefs as global carbon sinks-response to Greenhouse. Palaeogeography, Palaeoclimatology, Palaeoecology (2003). Available at: https://www.sciencedirect.com/science/article/abs/pii/003101829190172N. (Accessed: 18th February 2021)

[14] Peter Grinspoon, M. D. Cannabidiol (CBD) - what we know and what we don't. Harvard Health Blog (2020). Available at: https://www.health.harvard.edu/blog/cannabidiol-cbd-what-we-know-and-what-we-dont-2018082414476. (Accessed: 18th February 2021)

[15] Cannabidiol. National Center for Biotechnology Information. PubChem Compound Database Available at: https://pubchem.ncbi.nlm.nih.gov/compound/Cannabidiol. (Accessed: 18th February 2021)
[16] Cannabidiol. Wikipedia (2021). Available at: https://en.wikipedia.org/wiki/Cannabidiol. (Accessed: 18th February 2021)
[17] PottersWheel Available at: https://potterswheel.de/Pages/. (Accessed: 18th February 2021)
[18] PottersWheel. Wikipedia (2021). Available at: https://en.wikipedia.org/wiki/PottersWheel. (Accessed: 18th February 2021)
[19] Chemical process modeling. Wikipedia (2020). Available at: https://en.wikipedia.org/wiki/Chemical_process_modeling. (Accessed: 18th February 2021)