The Singularity Protocol

Florian Bansac*

Abstract

While many Artificial Intelligence researchers, companies and engineers are focusing their activities on the creation of Artificial General Intelligence with the hope of one day reaching the Technological Singularity, this paper takes a different approach. It presents a conceptual, speculative yet pragmatic organization of concepts to directly reach exponential scientific and technological progress. This paper is an attempt to design a process that automates the theories generation part of research and development through existing artificial intelligence algorithms and supporting technologies.

Notably relying upon recent developments in the conceptual and technical environment of graph neural networks, as well as upon multiple other technologies, the protocol breaks down the generation of scientific and technological theories in simple steps and propounds existing algorithm candidates to handle each step. With multiple scientific, mathematical and algorithmic references and sources for advanced readers, this paper nevertheless tries to make use of the simplest terms to facilitate the comprehension of a wider audience with minimal background in artificial intelligence.

Even though this paper describes a process that is still purely speculative for now, the Singularity Protocol does present a credible, structured and detailed approach to generate new scientific and technological theories at scale. And though it still needs to go through numerous adaptations, tests and computing challenges, this protocol is exclusively built upon existing technologies and it introduces a plan to gather and structure technical, financial and human resources so as to rapidly develop and implement an organization that could soon lead to the Technological Singularity.

Introduction

Background: Artificial General Intelligence and the Technological Singularity

The creation of Artificial General Intelligence - AGI - is expected to initiate an irrevocable change that will transform humanity beyond anything we can imagine. Though it is difficult to estimate what AGI could unravel beforehand, one of the main expectations is that AGI will be able to develop science and create new technologies much more rapidly than humans ever could. This event is usually referred to as the "Singularity", or more precisely the "Technological Singularity": the beginning of exponential scientific and technological progress.

Whether AGI can ever be created remains, however, to be demonstrated and done. Though important progress in the field of Artificial Intelligence - AI - has been made in recent years, thanks to the

* Correspondence: fbansac@ailephant.com 11 Sept 2018

availability of cheap computing power that now allows to process large amounts of data, the creation of AGI still remains a distant perspective for now.

Many hurdles, such as natural language understanding, consciousness, intentionality or even just properly defining and structuring human and machine intelligence themselves, need to be passed, or skirted, to build an AGI. And that is without even mentioning building an ethical AGI: one that doesn't launch the apocalyptic "Kill All Humans" scenario.

But, maybe, there could be another way: using <u>existing</u> AI technologies to attain the Singularity <u>without</u> first creating an Artificial General Intelligence.

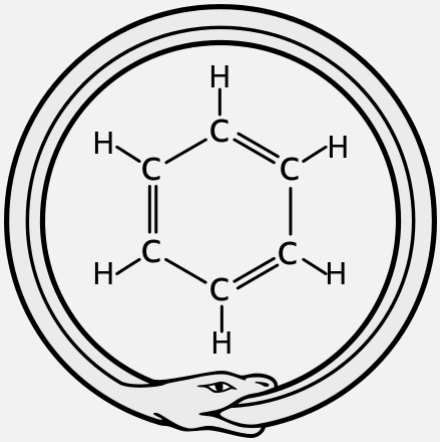
Writing with a lot of precautions and *ifs* and *coulds*, and counting on some major adaptation of existing technologies that are yet to be completed, the "protocol" (for lack of a better word) devised hereafter could put us on track towards exponential science and technology: the Singularity without the AGI, or maybe, before the AGI.

This protocol could be proved to be totally impractical because of the complexity of its source data, the necessary collaboration of people with conflicting interests or a number of other challenges. It could also be completely useless for now because of the enormous amounts of computational power it would require before any practical result can be reached. It may be quite obvious to many AI experts or even already being implemented in a similar pattern by certain AI companies. However, in any case, this protocol will at least have the merit of being written down in the present paper and hopefully spark a discussion, or help others with their AI experiments.

The Singularity Protocol principles and outline

Many artists, philosophers and scientists, including Albert Einstein who described how “combinatory play” is essential to productive thinking[1], have remarked that the creation of new ideas primarily, if not entirely, results from the combination of existing ideas, elements and concepts. New ideas, concepts, theories, etc. can be generated by associating and/or reorganizing existing ideas, sometimes from very diverse domains, in a new, original combination.

The entire protocol introduced here is thus built upon this central principle: **new scientific and technological progress *can* be created by combining existing science and technologies.**



The German organic chemist August Kekulé described the “Eureka” moment when he realized the structure of benzene:

“I was sitting, writing at my text-book; but the work did not progress; my thoughts were elsewhere. I turned my chair to the fire and dozed. Again the atoms were gamboling before my eyes. This time the smaller groups kept modestly in the background. My mental eye, rendered more acute by the repeated visions of the kind, could now distinguish larger structures of manifold conformation: long rows, sometimes more closely fitted together; all twining and twisting in snake-like motion. But look! What was that? One of the snakes had seized hold of its own tail, and the form whirled mockingly before my eyes. As if by a flash of lightning I awoke; and this time also I spent the rest of the night in working out the consequences of the hypothesis.”

- Extract from John Read “*From Alchemy to Chemistry*” (1957), see Wikipedia page on August Kekulé / Image: Haltopub

Figure 1. A famous example of combinatorial creativity

This protocol is an attempt to break down the theoretical part of the research and development – R&D – process into a sequence of simple operations so as to be processed by a machine. It proposes to match each operation with existing technologies and AI algorithms that appear to be good candidates to handle the task at hand. This protocol is also conceived to support a reinforcement learning system that aims at making it self-improving over time, thanks to a feedback loop that could be supported and enhanced by the Blockchain technology.

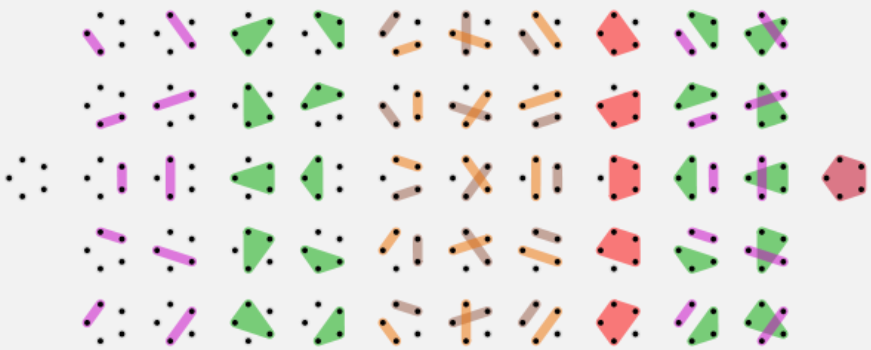
The protocol is articulated around **three key operational tasks**, corresponding to the three main sections of this paper:

1. Structure and label a graph database of existing scientific and technological advances detailing their specific features and utilities
2. Generate new theoretical advances by combining existing advances and/or their features from the database and prioritize these new theories for experimental testing
3. Collect feedback from experimental testing for recurrent improvement and reward feedback providers to incentivize the use of the protocol

The heart of the protocol thus lies in the generation of a vast number of theoretical new innovations (see Figure 2) and the ordering of these theories according to their desirability and/or practicality, which would be summed up in a utility score (or vector). The most interesting potential advances should then be tested through simulation and/or laboratory experiment to confirm or invalidate these theories and their actual utility.

The main advantage that the Singularity Protocol aims at providing is to drastically reduce the time spent on the conceptual part of the scientific research and technological engineering pipeline. As researchers and engineers go through a very time-consuming process of gathering thoughts to build new theories before experimental trials can begin (often leading nowhere), the automation of "combinatorial creativity", the combination of ideas and concepts to form new theories, could be greatly sped up and expanded upon thanks to computer algorithms.

By making computers do the heavy lifting, crossing many features and advances together in a profusion of combinations, before assessing the potential of each combination to keep only the most interesting for experimental verification, machines can play a big role in R&D. This "brute force" capacity for combinatorial creativity could lead us toward an exponential development of science and technology.



If we consider generating advances by combining only 5 existing features (without repeating them), we would already be able to generate 52 different combinations for unique, new potential advances.

Image: Watchduck

Figure 2. The combinatory "explosion"

Obviously, combining all sciences and technologies together is no small task. In order to initiate the process and obtain the first practical results faster, the protocol could, and most probably should, be made more effective by focusing on a reduced domain. In the beginning, it should also especially focus on areas of science and technologies where simulation is easier and lab tests are not mandatory.

If the ideas presented here can ever be successfully implemented in mathematical and computer sciences labs, or astrophysical simulators, it would be much easier in a subsequent stage to expand upon these results to other, more test-demanding fields, such as chemistry, biology, medicine, etc.

By reducing the domain upon which this protocol is tried, the initial testing and fine tuning phase should be much simpler than trying to embrace too vast and diverse subjects. Though this paper takes a general approach and ultimately aims at combining the most diverse fields, the initial scaling down of the scientific and technological domains to be considered will be all the more appropriate that this limitation will reduce the protocol's complexity, and therefore the resources required to develop it to rapidly reach a first success.

Though it is nowhere close to any success, it should be emphasized that the final aim of this protocol is to realize the expression of Humboldt and make "infinite use of finite means". The goal of aggregating, substituting and reorganizing scientific theories, technologies, features and concepts together is to obtain an ever-growing source of new technologies and grow the original database. Besides this self-sustaining database growth, the protocol's algorithms would also be programmed to be self-improving, so as to create better and better theories, faster and faster, and ultimately lead us towards the Technological Singularity.

Instead of developing an AGI hoping for the Singularity and not the Apocalypse, if an AGI can ever be created at all, we could rather (or in parallel) work on this protocol to directly aim at the Singularity. And though much of the existing technologies proposed here to handle the different operations still need some improvements and adaptations, and a lot of testing, the Singularity could be within our reach sooner if we can manage to make this protocol a reality.

Section 1: Pooling scientific and technological advances

Though science and technology can occasionally be used interchangeably, science is usually defined as the research of knowledge while technology is the application of scientific knowledge to solve problems and/or improve human life. According to these definitions, science is a preceding seed to reap the benefits of technology.

Since the goal of the protocol presented here is to achieve exponential scientific and technological progress, both science and technology will be considered as a single process going from scientific conceptual research to real-world technological applications. If this association is not abusive and that the protocol can be successfully implemented through most scientific and technological domains, it would generate new theories for both fundamental, theoretical science and new technological devices or engineering processes, ultimately leading to useful, practical applications in daily life.

Any initial implementation of the protocol will certainly start within a narrow domain of knowledge, as the complexity of the matter at hand and explosion of combinatorial possibilities will most probably present a challenge in terms of computing power requirements. To expand in the long term upon any success in the short term, obtaining a solid database of well-organized and well-labeled existing technologies will be an imperative foundation to build an effective AI model and train relevant Machine Learning - ML - algorithms.

1.1. Creating a graph database of scientific and technological advances

If the difference between science and technology is not always evident, the distinction between a *discovery* (the revealing of an already existing, but previously hidden or non-obvious mathematical or physical law, concept, rule, place, etc.), an *invention* (the creation of a new device or process) and an *innovation* (the improvement of a pre-existing device or process) is often even less clear when it comes to cataloging actual scientific and technological processes or devices.

In mathematics, for example, any new progress is as much an invention as it is a discovery: mathematicians searching for new theorems are imagining and building demonstrations as much as they are revealing existing numerical relations.

In daily life objects, the difference between inventions and innovations is not always clear either as features are added or replaced from previous objects to create new ones. What can clearly be defined as an invention compared to older technologies is often the result of adding or re-organizing a series of smaller, more recent innovations. The invention of autonomous cars is often considered as a breakthrough that will greatly impact human life, but self-driving cars are also just a feature improvement from normal cars. So can smartphones be considered as improvements of regular cell phones (with the intermediate stage of the Blackberry) or the internet be considered as computers linked together just as people were connected together before with land telephone lines.

Since any effective computer system, especially a set of ML algorithms, needs to be built upon certain formal definitions that can be applied in most cases, the model presented here will therefore rely upon the central notion of a scientific or technological **advance**. In this conceptual setting, an advance is defined as a set of **features**, in specific **relations** to each other, and from which a global **utility** results from the benefits the advance provides. The utility of an advance is thus a function of the arrangement of its features in their specific relations to each other (or absence of relations).

In order to place ourselves in a convenient and structured representation of science and technology that can be handled by machines, advances can be considered as **graphs** of features in relations to each other, from which a global utility attribute is derived. Thankfully, that form of *knowledge graph* fits very well in the "**Graph Networks**" - GN - conceptual and technical environment summarized in the fundamental paper on *"Relational inductive biases, deep learning, and graph networks"*, published in June 2018 by a group of some of the world's most prominent AI researchers[2]. This central paper (hereafter mentioned as *"GN paper*" to facilitate reading) regroups and consolidates numerous sources into a standardized model from which this protocol heavily borrows.

Relying upon GN is an especially practical model for advances in this protocol since GN are structured representations that support arbitrary relational reasoning. This characteristic allows GN to be used in a number of ways to build the protocol as GN have demonstrated that they can "support combinatorial generalization"[3], "generalize well to problems of much different sizes than they have been trained upon"[4] and have already "been explored in a diverse range of problem domains, across supervised, semi-supervised, unsupervised and reinforcement - RL - learning settings"[5].

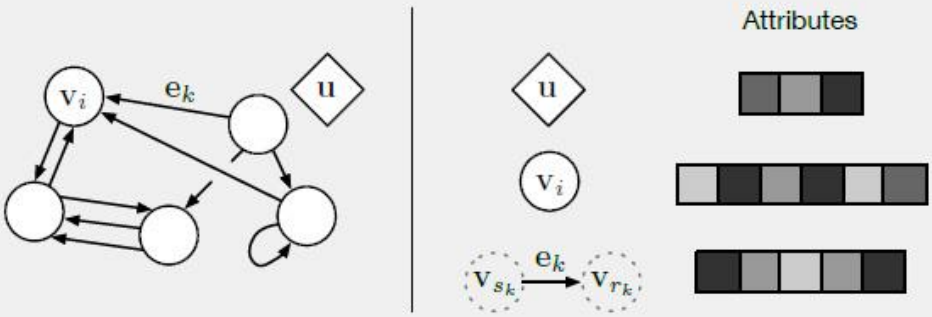
Integrating further into the GN paper's standardized approach (see Figure 3), any advance in the protocol is assimilated to a **graph**, which are conceived as a **3-tuple G = (u, *V*, *E*)**, where:

- **u** is ***utility*** of the advance G - assimilated to the **global attribute** in the GN paper
- ***V*** is the ***set of features*** of the advance G - assimilated to the **set of nodes** in the GN paper: $V = \{v_i\}$, with i=1:N^v, where N^v is the cardinality of V
- ***E*** is the ***set of relations*** between these features - assimilated to the **set of edges** in the GN paper: $E = \{(e_k, r_k, s_k)\}$, with k=1:N^e, where N^e is the cardinality of E

With this formal definition, the GN paper also describes the **attributes** of graphs: characteristics of the global attribute (utility), nodes (features) and edges (relations) which can be encoded as vectors, sets or other graphs. The diagram and explanations of Figure 3 reproduce these details here so as to facilitate

comprehension and enhance practicality for the reader. Please refer to the GN paper for complete descriptions of the graph formalism, explanations on graph networks and guidance for computations.

The following sections 1.2 and 1.3 detail the specific characteristics of advances graphs that will be critical to train generative ML algorithms to combine existing advances, features and relations so as to create new potential scientific and technological advances. Building, or obtaining, a database of advances defined as graphs will therefore be a key starting point to implement the subsequent operations of the protocol.



Here we use "graph" to mean a directed, attributed multi-graph with a global attribute. In our terminology, a node is denoted as $\mathbf{v}_i$, an edge as $\mathbf{e}_k$, and the global attributes as $\mathbf{u}$. We also use s_k and r_k to indicate the indices of the sender and receiver nodes (see below), respectively, for edge k. To be more precise, we define these terms as:

Directed : one-way edges, from a "sender" node to a "receiver" node.
Attribute : properties that can be encoded as a vector, set, or even another graph.
Attributed : edges and vertices have attributes associated with them.
Global attribute : a graph-level attribute.
Multi-graph : there can be more than one edge between vertices, including self-edges.

Figure 3. Definition of a "graph"
Reproduced from the paper *"Relational inductive biases, deep learning, and graph networks"*[6]

Since a general graph database of all existing sciences and technologies may not be readily available to the public, users of the protocol presented here could select a limited set of advances (and maybe add proprietary technologies, patents, expertise or products) that they want to consider in order to stay most relevant to their own R&D efforts, so as to reduce and focus the initial fastidious work of building a graph database (which could also be proprietary). The narrower the domain of consideration, the easier and faster it will be to constitute a usable database of advances graphs with their related features and utility, and the more effective the protocol may effectively work in the short term.

To regroup individualized efforts for the wider public good, a system of reward for contributions to assemble various databases from different sources will also be sketched in section 3 in order to consolidate a single, open graph database at a later stage of the protocol development.

1.2. Defining features and relations of features

In the model of advances graphs, be it an actual material object (vehicles, electric components, molecules, etc.) or immaterial concept (time period, physical properties, weather patterns, software agents, etc.), **features** are the different components of an advance.

Features are graphically represented as **nodes** of the advance graph: they are linked together as sources and/or receptors of relations vertices to other features, and they can be attributed. All the features of an advance are functionally expressed as the set of nodes *V* presented in the previous part 1.1. Each node is given an attribute which can be encoded as a vector, set or graph, and may especially represent a sequence of text corresponding to words or sentences, which may stand for a definition, label, characteristic or property of a material or immaterial feature (see example in Figure 4).

Remark that the level of precision of the graph of a single same advance may differ according to the needs and references of specific protocol users. If a user of the protocol has to create their own proprietary database according to their actual R&D needs and focus, features that are out of scope could optionally be grouped together in summarizing clusters to simplify the process. A car manufacturer could break down its cars and technologies into engine parts, materials and chemical reactions, while a hardware producing company would go down to CPUs, integrated circuits and transistors, and a fundamental physics research lab would ultimately rely upon particles and electromagnetic forces.

The most precise graphs that would ultimately compose the open, general graph database for the benefits of the larger community of users should be detailed down to the most basic sets of concepts and primordial building blocks of any technology: fundamental mechanical rules and properties, materials involved, environmental conditions, size, weight, or other relevant concepts (such as parity/imparity, bottlenecks, symmetry, geometric shapes, mirror images, series or parallel circuits, alternative or continuous currents, failsafe mechanisms…). Standardization of graphs collected from different users and the possibility to consider simpler versions will therefore be critical in the consolidation of the protocol's main database and its algorithms in the long run, as sketched in section 3.

In the formal model of advances graphs, the **relations** between the features are simple interactions between two features that can either be expressed as an action in a sequential pattern (such as action/reaction, cause/consequence, transmission of an information, electric signal, magnetic or gravitational attraction, etc.) or a spatial associations between two features (association, inclusion/exclusion, relative position, correlation between numerical data, scale, scope, bond between atoms, etc.).

Relations can be represented with an **oriented vertex** linking a source node and a receiver node that can be attributed. They are functionally expressed as the set of edges *E* presented in the previous part 1.1. Each vertex is given an attribute which can be encoded as a vector, set or graph, and may especially represent a sequence of text corresponding to words or sentences, which may stand for a definition or label of any temporal, spatial or other arbitrary interaction (see example in Figure 4).

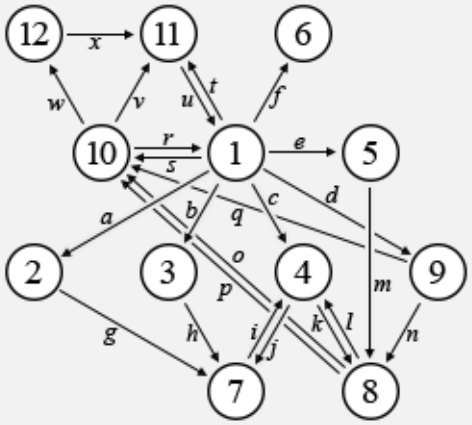
Since it can be assumed that at any given time there is a **finite number** of existing scientific and technological advances, there is also necessarily a finite number of features and a finite number of relations between those features. This characteristic would suggest that the creation and treatment of such a graph database with finite elements is potentially feasible by a machine given enough computing power and proper constraints in the combinatory associations that can be operated.

Obtaining or building a graph database of advances is thus an imperative prerequisite to constitute a pool of features, advances and relations. This database will be used to train and feed the generative combinatorial algorithms that will mix and match advances and/or features and/or relations to create new potential advance in the following stages. To make sure these generated advances are practical, the original advances graphs in the initial database will also have to be characterized by their utility, a key characteristic upon which the protocol will rely, as explained in the following section 1.3.

Advance: Autonomous Car (Simplified)

Utility (global attribute): u
Category: Vehicle
Simple definition: Optionally Autonomous, 4-Wheeled Passenger Vehicle
Status: Production - For sale to general public
Characteristics:

Main benefits:	*Main disadvantages:*	*Other characteristics:*
Max speed: 180 km / h	Price: 30,000 $	Fuel conso.: 1 liter / 100 km
Passengers cap.: 5 adults	Fuel costs: 1.50 $ / liter	Avg. lifetime: 200,000 km
Trunk cap.: 300 kg	CO_2 emission: 1 kg / 100 km	Requir.: Driver licence
Self-driving: Optional	Parking space: 10 sqm	Periodic tech. control



Features (nodes): v_i
1. driver
2. steering wheel
3. brakes
4. gears system
5. accelerator
6. signals & headlights
7. wheels
8. engine & battery
9. fuel tank
10. dashboard/computer
11. auto-pilot software
12. environment sensors

Relations (vertices): e_k

a. driver turns steering wheel left/right
b. driver gradually pushes brakes more/less
c. driver switches gears up/down
d. driver periodically refills fuel tank
e. driver gradually pushes accelerator more/less
f. driver switches signals & headlights on/off
g. steering wheel orients front wheels left/right
h. brakes decelarate wheels
i. wheels transmits decelaration to gears system
j. gears system transmits accelaration to wheels
k. gears system transmits decelaration to engine
l. engine transmits accelaration to gears system
m. accelerator injects more/less fuel to engine
n. fuel tank provides fuel to engine
o. engine/battery powers computer
p. engine/battery informs status to computer/dashboard
q. fuel tank informs status to computer/dashboard
r. computer/dashboard informs all statuses to driver
s. driver switches computer functions on/off
t. driver switches auto-pilot on/off
u. auto-pilot replaces driver for *a,b,c,e,f*
v. computer powers auto-pilot software
w. computer powers environment sensors
x. environment sensors inform conditions to auto-pilot

Figure 4. An autonomous car defined as a simple graph
(sequential relations and mechanical features only)

1.3. Defining the utility of each advance

The goal of formalizing scientific and technological advances in a graph database is to use this data to train ML algorithms to automatically generate a large number of new potential advances. Since many of these theories will assuredly be useless, if not absurd (yet maybe still interesting to keep for subsequent recombination), and that, within the fraction of interesting ones, some will be more useful

than others, it also seems critical to estimate the viability of these generated advances. To do so, a score (or vector) should be associated to each generated theory to prioritize them for experimental testing.

The utility of scientific and technological advances therefore seems to be a convenient criteria to regroup positive, negative or other characteristics, and compare advances together. Advances provide more or less benefits (speed, power, convenience, comfort, ease of use, beauty, taste, etc.) and/or decreases resources required to produce or use them (time, costs, effort, pollution, etc.), and/or their feasibility and usefulness according to current means available. Two advances can then be compared through these sets of characteristics, at least in the dimensions where they both can be attributed a score.

Even though defining the complete **utility vector** of positive, negative and neutral characteristics for each and every advance is no trivial task (especially if intangible concepts such as beauty, pleasure or pain are considered), it would also seem that it is feasible, at least for limited number of related advances. Many characteristics can be quantified easily, and for many advances, these characteristics are readily available in the form of technical product specifications (from manufacturers) or through the definition and properties of scientific theories and concepts.

Measuring the **absolute utility** of each and every advance ever discovered, invented or innovated is a daunting task, but it is clearly more feasible to compare a sub-set of related advances and their relative utilities. Comparing a limited number of advances to previous advances they replace or related advances they complement is also much more straightforward, especially if only certain specific utility characteristics are considered.

Comparing the utility of a self-driving car to complex numbers is difficult and probably of little interest, but comparing it to a regular car can be much easier and useful if we place ourselves in the domain of automotive transport. Compared to a regular car, a self-driving car provides x % more "free" time to its driver and an electric car reduces fuel spending of n dollars and pollution of m kg of CO_2. Limiting the domain of concern is evidently also more interesting for R&D laboratories, as it will help them focus on their specific area of expertise and target value-creating innovations.

The **relative utility** vectors should therefore only be precise enough, at first, to consider the dimensions of interest to each specific R&D lab, researcher or engineer. Though comparing broader sets of utility dimensions will be most interesting in a later stage, pinpointing limited benefits will be critical in the initial stage to adjust algorithms for maximal efficiency.

Re calibrating relative utility vectors between multiple advances should be relatively simple at a later stage to consolidate absolute utilities. If the process of defining relative utilities detailed here does provide any success in limited, specific domains, regrouping the databases from various users and domains should be feasible to consolidate a more general, open graph database with more absolute comparisons and rankings of utilities (more on this consolidation in section 3).

Once again, the consolidation of absolute utilities in the aggregated database should provide significant benefits in the long run. Generating graphs from the most diverse domains with diverse features and relations should help the protocol's algorithms create the most serendipitous theories in the long run: graphs of advances that scientists and engineers could not have imagined as they would be composed of the most unintuitive features and relations associations, and yet that are still valid and provide a utility.

Obviously, the generative and prioritizing algorithms sketched in section 2 will only provide any benefit if they can sort the most useful theories with meaningful utility vectors. Properly defining the utility of existing advances will thus be one of the most critical factor to train ML algorithms to estimate the potential utility of future generated advances. In a subsequent stage, the utility of advances will also provide a source of measurement for a RL process of the protocol's algorithms: it will be considered as the *reward* to be maximized when updating the generative algorithms (see section 3).

1.4. Automating the creation of the graph database

It would seem that a global graph database of scientific and technological advances defined in the required format of 3-tuple graphs does not exist yet, or is not publicly available, especially if we consider adding utility vectors. However, some open databases are available to be processed into the desired graph database format, such as Wikidata (from Wikipedia) or related databases like Yago and DBPedia, to which other public and/or private resources, patents, research papers could also be associated in secondary (maybe proprietary) databases. Some fine-tuning will be necessary but the construction of such a database seems feasible, provided the utility vectors can be added, retrieved from manufacturers' technical specifications or from textual definitions of theorems and scientific properties.

With the perfect method still to be thoroughly defined, it is however not a far stretch of the mind that some existing feature learning algorithms could also be adapted with text classification, to parse through a database of advances and summarize scientific concepts and technologies in features and relations of features. Several existing algorithms could be adapted to drive the automating feature extraction, notably Recurrent Neural Networks which have shown great success and versatility working with sequences (such as sentences), and other algorithms transforming text or structures to vectors, such as word2vec[7] or structure2vec[8].

With regards to automating the association of a utility vector to each advance, certain graph networks do present characteristics that could also ease the process in a second step. Message Passing Neural Networks[9], especially, seem to be great candidates to infer a global attribute from graphs, such as the utility, once these graphs have been created (see section 2.3).Anyway, if an automated method cannot be readily implemented to obtain the necessary graph database, the fastidious and time-consuming method of building it manually can always be conducted by combing through the available resources to specify the graph of each advance with its inter-related features, and add the utility vector corresponding to its characteristics. However, an automated method should be possible to implement: after all, AI is already beating humans at reading, maybe it could just as well summarize sciences and technologies in features and relations, and associate their utility vectors.

Section 2: Generating useful theoretical advances

Research and development is a time consuming process that often leads nowhere. Ideas take a lot of time to be properly structured into scientific theories or engineering processes before they can be tested through judiciously designed experiments that will prove them right or wrong. If theories are proven wrong, which happens most of the time, researchers, scientists or engineers have to find errors to formulate solutions or start the entire process all over again.

The present protocol primarily aims at handling this "theory building" part to speed up and expand the scope of research and development. Its goal is to generate theories for scientific and technological advances faster and in a larger number than can ever be done by scientists and engineers, virtually shrinking the theories forming process from weeks, months or even years down to hours or minutes.

By automatically combining (substituting, adding, removing and reorganizing) advances and/or features and/or the relations between these features from the database of advances described in section 1, new scientific and technological processes and devices would be conceived by machines faster than current, human-managed laboratories and plants ever could verify.

Though generating potential advances faster and in greater number would save time in theory building, researchers and engineers could soon be overwhelmed by the number of theories to experimentally verify. Avoiding the creation of absurd theories and prioritizing the interesting ones will therefore be critical to successfully implement the protocol and reach the real, practical and useful scientific and technological advances that will drive us towards the Singularity.

2.1. Mapping valid associations of advances, features and relations

The essence of this protocol is the automation of combinatorial creativity applied to science and technology. Even though it is a simple principle to understand, combining and recombining advances and features is a much more difficult problem when it comes to programming a machine to do so. Some ML algorithms have performed well at generating new data for some time, but we are only now seeing the first successful attempts at generating graphs, and generating graphs with arbitrary relations and nodes is still in its infancy.

Current approaches show that a critical first step is to develop an "understanding" of how graphs can or cannot work. In the protocol, that would mean discerning how certain types of features can interact or be isolated from each other, and what type of relations these features can entertain (or not). That is one of the main uses of the database introduced in section 1: providing example data on scientific and technological advances formalized as graphs to be parsed by ML algorithms so that they can extract a map of how certain features and relations can be clustered together or not.

Following the terminology employed in the GN paper, this **Relational Inductive Bias** - RIB - must be extracted from the database to create a latent graph (a *kernel*) of the underlying spatial and/or sequential clusters of features and relations, so that it can be propagated as a model or used as a constraint in the following generative phase. This latent graph of valid "node to node" structures will be integrated in the architecture of the generative algorithms, introduced in section 2.2, to either prevent them from creating absurd advances, *a priori*, or to filter absurd advances out, *a posteriori*.

A key aspect to consider when selecting the best use of the extracted RIB during the generation of advances is the management of the computing resources available, and therefore the RIB extraction. Since the algorithms will use a large graph database to generate a potentially larger set of new graphs, it will be critical for the protocol users to plan in advance how they will extract and use the RIB to either minimize the resources needed to obtain acceptable results, or to reach the very best potential advances provided more computing resources are available.

Minimizing resources to obtain acceptable results in the relatively limited domain of molecules, a recent research project[10] has used *junction trees* to separate graphs into sub-graphs and determine cluster of valid spatial relations between atoms. With this form of RIB, using these extracted valid clusters as building blocks in a subsequent generative algorithm, this model has built exclusively valid molecules whereas several previous sequential methods generated multiple incoherent molecules with invalid atomic bonds from random associations.

With the opposite aim of reaching the very best results if resources are not scarce, the GN paper also presents some graph-focused algorithms that could be efficient at discerning relational inductive biases, maybe, at least, in some specific domains and situations (see Figure 7). These especially include, *message-passing neural networks*[11] which are able to convey rules and constraints (that could also be used to estimate the utility vector as explained in section 2.3), or other graph networks[12].

Both of these methods will be presented in more details in section 2.2, the generative part of the algorithms. Again, provided sufficient computing power is available, and depending on the domain(s) considered, multiple algorithms could also be used sequentially or in a parallel to extract the RIB latent graph, the "map" or "catalog" of valid advances, features and interactions of features to be used in the following generative step.

2.2. Combining existing advances, features and relations to generate new advances

Several models have been proposed to successfully generate tables of data, images, sounds or text for some time, but structured representations such as knowledge graphs bring more complex problems of coherence that only now start to be properly addressed. If graphs were to be constructed by aggregating

randomly selected nodes, most of the graphs created wouldn't make any sense. That's why the **relational inductive biases** as well as the **combinatorial parameters** such as the pool of advances, features and relations to consider and the maximum depth of combinations (number of nodes and vertices that can be combined together, and what sets of advances can be drawn from) must be taken into account in the generative process.

A recent architecture[13] has successfully been implemented to generate graphs through a sequential process, defining what could be considered a "standard method" of adding node by node and linking them with vertices according to a probability score (see figure 5). Compounding from a previously extracted catalog of features and relations, this *probabilistic algorithm* builds advances graphs by making decisions of adding features (or not) and finding relevant relations between them. The algorithm can support some parameters (such as building from a given feature or sub-graph, using RIB to limit features associations, build only from a set of features and infer relations from the RIB kernel instead of combining existing features, limiting depth of combinations…) and its results can be discriminated, through a conditional probability of relevance.

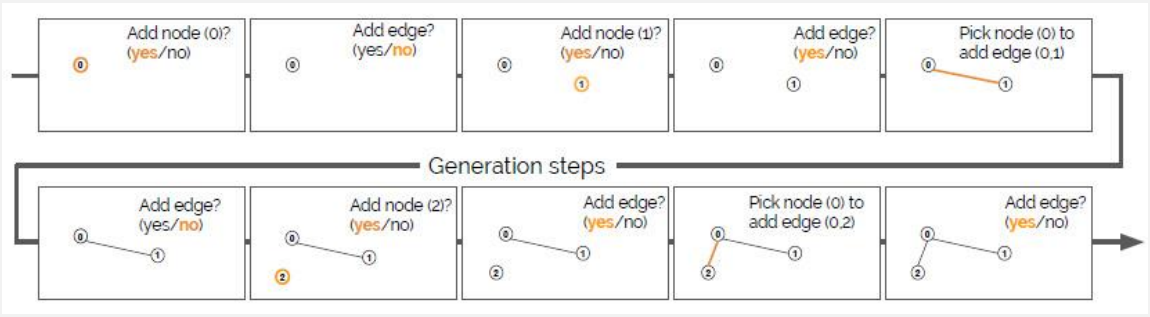


Figure 5. Depiction of the steps taken during the generative process of a graph
Reproduced from the paper *"Learning Deep Generative Models of Graphs"*[14]

The conditional parameters and the RIB will also help to tackle another major challenge that arises when it comes to the actual generation of new advances graphs, given the combinatorial explosion of possibilities and complexity that grow quadratically with the source database. Any generative algorithm associating these features and relations will probably pose the problem of the computing power required, and therefore the resources available to individual users on their own machines or renting needs. Though computing power is now relatively cheap and available, the combinatorial creation of large numbers of graphs from large databases still may cause serious computing challenges.

As previously mentioned, this "standard" architecture can be enhance through two methods, presented hereafter, depending on the availability of a user's computing resources. In the first method, the RIB is used "strictly", as a set of **constraints** for conditional generation of graphs within a limited domain and to minimize the resources required to obtain acceptable results. In the second method, the RIB is used "loosely", as a **filter** to bring the most interesting advances up from a large pool of generated graphs, and push down the ones that do not correspond to the characteristics of real-world advances. Both methods use the RIB, the first for *belief propagation* of valid criteria into the generative algorithm, the second as a *discriminator* of validity/invalidity after the generative algorithm (see figure 6).

Enhancement method 1: increasing relevance by generating advances from valid features clusters

The first method makes a rigorous use of the relational inductive bias. By propagating the RIB criteria of validity extracted in section 2.1 into the generative algorithm, a minimal number of advances is

created, ensuring all of them correspond to criteria of validity. This method implies that the RIB criteria must be rigorously defined, and therefore must correspond to a limited set of advances and/or domain, from which all the validity characteristics must be extracted in the RIB kernel.

In a limited and well-structured domain, this approach has already proved that it can generate useful results. Considering molecules for the pharmaceutical industry, this method has led to the generation of more useful molecules than other methods, especially demonstrating that all of the molecules generated were valid whereas other models generated large proportion of invalid (useless) molecules.

The results from an application to pharmaceuticals[15] highlights that the generation of graphs by **assembling only valid clusters** of molecules (valid sub-graphs), that were first extracted into junction trees, ensures that all generated new molecules are fundamentally valid molecules. Since this experiment uses a *Variational Autoencoder* - VAE - (an algorithm that generates results fitting certain characteristics, especially the previously extracted RIB kernel) and Bayesian optimization (a process to select the output of the algorithm according to certain desired characteristics), not only does this method allow to generate results that exclusively fit the RIB criteria, it also allows to arbitrarily choose certain specific features (and their properties) to serve as a base for the generative algorithm to build upon.

It would therefore seem that this first method could be successful in some very specific, narrow and spatially deterministic domains, somehow similar molecule generation, where RIB can be rigorously used as a set of constraints *a priori*. Some domains with similar characteristics, where it may be possible to obtain new theoretical advances through this method, could include the design of electric circuits and computer chips and the design of mechanical devices and engines.

This method may also be useful for users who wish to focus their R&D efforts around new applications of a very specific, preexisting expertise or product. Since certain properties or characteristics can be selected to generate new advances, a user could choose to only generate new advances that are built by composing around the pre-defined graphs of their expertise or product. A manufacturer of specific electrical or chemical components could thus focus on finding new applications that use these specific components.

A major downfall of this method is that using *rigorously* the RIB previously extracted from the graph database, the algorithm can exclusively generate advances that reproduce structural associations of nodes and relations that are already explicitly mentioned in the database. Avoiding any other type of relation inherently limits the capacity for novelty of this generative algorithm, and though it may save a lot of time and computing power by eliminating entire branches of useless advances, this method also will structurally miss the occasional original association that could work and bring the most innovative benefits.

Enhancement method 2: increasing novelty by generating numerous advances and sorting the most promising ones

The second method relies on a “brute force” approach that particularly makes more use of the machines potential. It is inspired by *Generative Adversarial Networks* - GAN -, which opposes two ML algorithms, one *generator* creates data and one *discriminator* assesses if this new data corresponds to a set of valid test data, and has already been adapted to graph structures[16]. Here, assembling features and relations from a large graph database (see section 1) of diverse advances from diverse domains, a generative algorithm would create all the possible advances graphs combinations.

The generative algorithm could include certain restrictions, such as a probability of coherence estimated from the RIB to limit the obvious absurdities, but the goal is to make associations in greater number and with more diverse sources so as to generate graphs that produce more diverse combinations (the number of nodes and/or vertices may also be capped to limit computing power needs). This approach aims at

reaching the most groundbreaking advances by **making associations of the most diverse features and relations** that would otherwise be too "unobvious" for researchers and engineers to conceive.

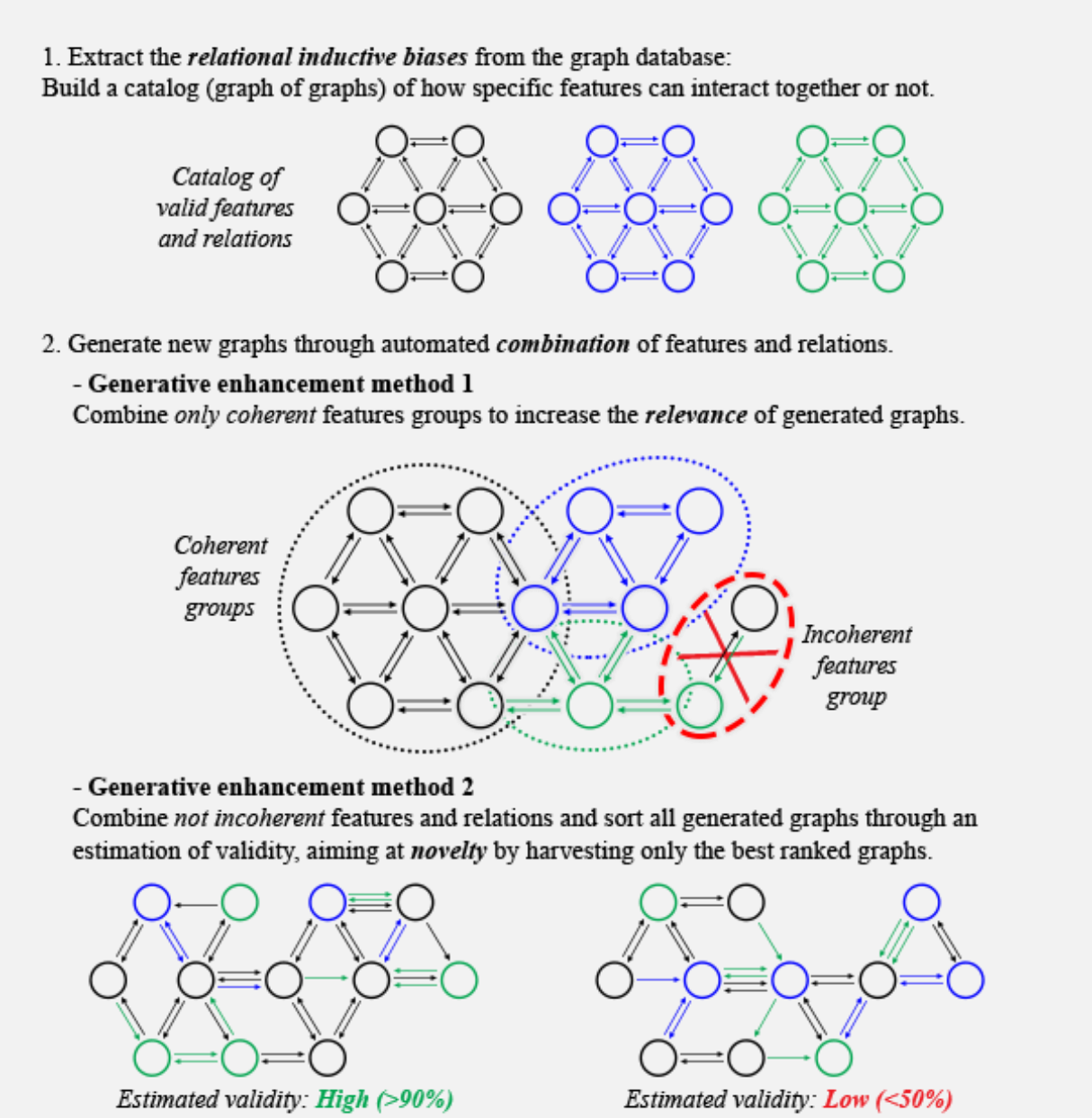


Figure 6. Generating new advances graphs and consideration of *relational inductive bias*

Following this generative first step, a discriminative second step would then rank advances according to an estimated validity score (which could fused with the Utility Estimator of section 2.3). Assessing generated advances "loosely" against the RIB criteria, they would be ranked according to the number of probable/improbable validity of local associations of features and relations. The goal of that second step is to **rank the most promising advances first** and push down the least promising ones. Instead of being a valid/invalid criteria, the RIB is here a *softmax classifier*, which makes an order between the graphs rather than a true/false test.

Obviously, this method would lead to the creation of numerous invalid (and absurd) advances graphs. This method would therefore make a rather inefficient use of computing resources, which would be all

the more used to no end that the database of advances to combine from is large. Nevertheless, the aim here is to extract a few useful nuggets from a mass of useless material, hoping that the benefits obtained from these very few, hopefully more original ideas may counterweight the costs of wasted computing power and time needed to reach them.

Other generative methods

Choosing to use one method or the other will actually depend on the goals and available resources of the individual users of the protocol's algorithms: either obtaining acceptable, focused innovations with limited resources or aiming at more original breakthroughs with more resources to spare. Other solutions to generate advances could also come from the **combination of both methods**. GAN and VAE models have already been combined together to generate images[17], maybe they could also be adapted together for graphs generation. Various other graph networks architectures[18] could also be used for the generative algorithms as well as for the extraction of the RIB.

Besides these potential models, future tests may prove that other algorithms that still need to be created or adapted from other sources could be more effective. Depending on domains, sub-sets of advances and features, and the types of spatial and/or sequential relations considered, various ML operations may be required to extract the RIB from the database, combine the actual features and relations and assess the validity and utility of results.

Hence, given the profusion of possible candidates, an **overarching model-selecting or model-assembling algorithm**, relying upon Bayesian conditional probabilities or an evolutionary approach[19], could be the best architecture to select, mix and match the most effective algorithms. It would process by associating and integrating various combinations of Graph Networks with plain combinatory algorithms and/or the arsenal of more classic ML algorithms, such as hidden, convolutional and recurrent layers and measure the efficiency of each architecture for each domain and/or sub-set of sciences and technologies to monitor and improve the generative process in the long run through RL updates (sketched in section 3).

This complex study goes beyond the scope of this paper, but even though the present protocol is still purely speculative for now, whatever future tests may conclude, we do already possess several strong leads to automate the generation of scientific and technological advances graphs. If some of these existing algorithms can be adapted to generate graphs with features and relations relying upon arbitrary characteristics, and estimate their potential utility (presented in section 2.3), the protocol will set us well on the way towards the Singularity.

2.3. Estimating the utility of generated graphs to prioritize advances candidates

After the combinatorial algorithms generate advances graphs, these candidate theories still need to be tested in laboratories to be validated. Since the number of new generated graphs should eventually be way too large to be tested by any pool of researchers, engineers and R&D labs, and that an important part of these generated theories will inevitably be impractical, if not entirely absurd, a second system should be included as a final step in the generation process to sort the generated graphs by their potential, according to the utility score introduced in section 1.3.

This step may be redundant with the discriminator presented in the second generative enhancement method (of the previous section 2.2), if the validity and utility can be fused together through a single discriminator. However, it is nonetheless presented separately here because this utility discriminator would need to be applied to all generated graphs, whatever the generative method may be, "standard" probabilistic method, VAE or GAN inspired enhancements (method 1 and 2) or other methods.

When a combinatorial algorithm generates a new theoretical advance, a corresponding utility vector should be associated with each new graph by estimating the advance's new benefits and/or

reductions in resources required, if they cannot be calculated precisely. All the generated graphs can then be ranked according to their utility vector, which will serve as a **discriminative classifier** to prioritize the generated advances for experimental verification via simulation and/or lab tests, depending on the scientific and technological domain and/or type of advance.

Developing the utility estimating algorithm seems not to be an exceptionally complex task, if preexisting advances graphs and their utilities can be appropriately defined, as sketched in section 1.3. *Multilayer Perceptrons* and other supervised learning algorithms are already successful at accomplishing similar tasks, provided the database upon which they are trained is properly built and labelled. However, since we are evolving in the conceptual environment of graphs, relying upon the Graph Network Blocks[20] - GN Blocks -,"graph-to-graph modules", and, in particular, Message-Passing Neural Networks - MPNN - would seem more appropriate.

The property of GN Blocks and MPNN to pass rules and constraints in order to compile a global attribute (which can be a vector) fits well into the protocol's requirements to compute the estimated utility of newly generated advances graphs, which may include their practical feasibility with current means and technologies as one of the vector's coordinates. Figure 7 is a reproduction of these two "graph-to-graph" configurations from the GN paper[21] which also recapitulates various other architectures to compute the global attribute, through the update of a graph.

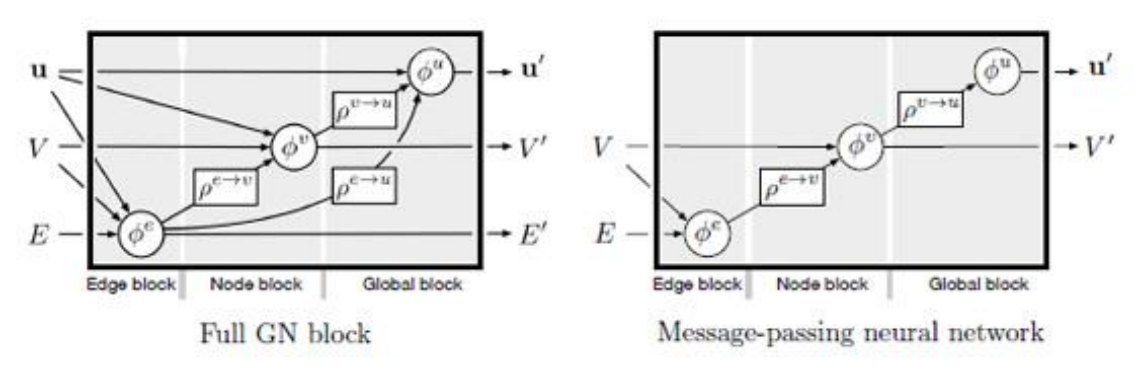


Figure 7. Two "graph-to-graph" configurations candidates to estimate the *utility* of generated advance.
Graph reproduced, with the following legend from the paper
"Relational inductive biases, deep learning, and graph networks"[22]
A Full GN Block predicts node, edge, and global output attributes based on incoming node, edge and global attributes.
A Message-passing neural network predicts node and global output attributes based on incoming node and edge attributes.

The standard GN Block configuration, as well as the MPNN, are strong candidates to handle the protocol's utility vector extraction if we consider a graph, G = *(E, V, u)*, such as the extracted RIB graph kernel or an input existing advance graph, and a generated advance graph G' = *(E', V', u')*. *E'*, the new advances' features with their relations, *V'*, are obtained through the generative combination, introduced in section 2.2, of the original features catalog from the graph database presented in section 1, under the conditions of the RIB learned from section 2.1. In that configuration *u'* would then be the utility vector calculated by a judiciously defined GN Block, or rather, an arrangement of different GN Blocks into a "GN core"[23], multiple consecutive graph network operations.

With such fitting candidate algorithms to calculate the estimated utility of new advances, some form of *softmax function* would be appropriate to rank the generated graphs according to their estimated utility

vectors. It would then be in the hand of the protocol's users, R&D labs, researchers and engineers, to pick their most interesting theories from the top of the softmax-ordered list to design and conduct the experimental validation (or disproval) of the generated graphs and measure the actual utility vector.

The combinatorial and prioritizing algorithms are presented here as separate steps in the protocol for more clarity, but compounding these two algorithms in a single GN Block, or GN Core, or through other algorithmic architectures may be a better solution. Having resorted to this clarification to ensure the readers' comprehension of the algorithmic operation to be developed,

1. pooling of the advances graphs, features and relations,
2. generation of new graphs from this pool,
3. prioritizing these generated graphs through their estimated utility,

section 3 structures ideas and technological means to consolidate the protocol through feedback collection and grow its community of users by rewarding them for their collaboration.

Section 3: Consolidating the protocol

With the core of the protocol's algorithms and training data clarified, the last section intends to make the system self-sustaining and self-improving. The essential source of this virtuous circle lies in the collection of feedback from experimental verification of the generated advances graphs to fuel the growth the community of users in addition to improving the algorithms and expanding the underlying database.

Though various versions of the algorithms may be used by specific users to tackle their respective domains of research, the protocol's objective is to be consolidated in a single system, or a very limited number of coordinated systems. The end goal of this consolidated protocol it to reach the most promising technologies faster by combining advances and features from more and more diverse domains of science and technology.

Not only does a growing database provide more features and relations to combine, the algorithms consolidation process also aims at taking advantage of **transfer learning**, to provoke cross-domain "serendipity". The algorithms leading to experimentally verified generated graphs should then be reused, or receive increased weight, to generate new graphs in other domains that present similar situations.

Yet, this consolidation also relies upon the willing participation from various public and private actors that could be competitors. To stimulate the contribution of potentially rival organizations, and avoiding a naïve utopian view of uninterested collaboration, a system of financial retributions should also be conceived to reward feedback providers. Since a system of encryption must be used to communicate experimental metadata while ensuring privacy, the protocol proposes to make use of **smart contracts** and an underlying **cryptocurrency** to serve both purposes: securing privacy and data received from feedback providers while supporting a corresponding financial retribution system.

3.1. Experimentally verifying advances candidates and their utility to drive reinforcement learning

After obtaining a list of generated graphs of potential advances ordered by estimated utility, an experimental process must confirm the validity and the actual utility of advances. Since it takes care of theory formation, the main benefit of the protocol is thus to help engineers and researchers channel their efforts on experiment design or simulation, who then prove or disprove the most interesting theories.

In the long run, that experimental task could be managed by fully automated labs, and we see some developments in that direction with the likes of the Emerald Cloud Lab already conducting experiences

for third parties. For now however, even though the protocol would relieve R&D labs, scientists and engineers from a heavy, conceptual effort, they will still have much experimental work ahead, especially if the protocol does lead toward the exponential generation of new theories.

The experimental verification (or invalidation) of generated graphs together with the measurement of the discrepancy between their actual utility confirmed by engineers and scientists and the estimated utility from the algorithm finally would also provide a convenient source of feedback to improve the protocol. The experimental validation and resulting utility discrepancy can be treated as a series of quantifiable factors of improvements or corrections: **reward** signals to eliminate misleading algorithmic processes and improve the combinatorial and utility-estimating algorithms over time through a reinforced learning approach.

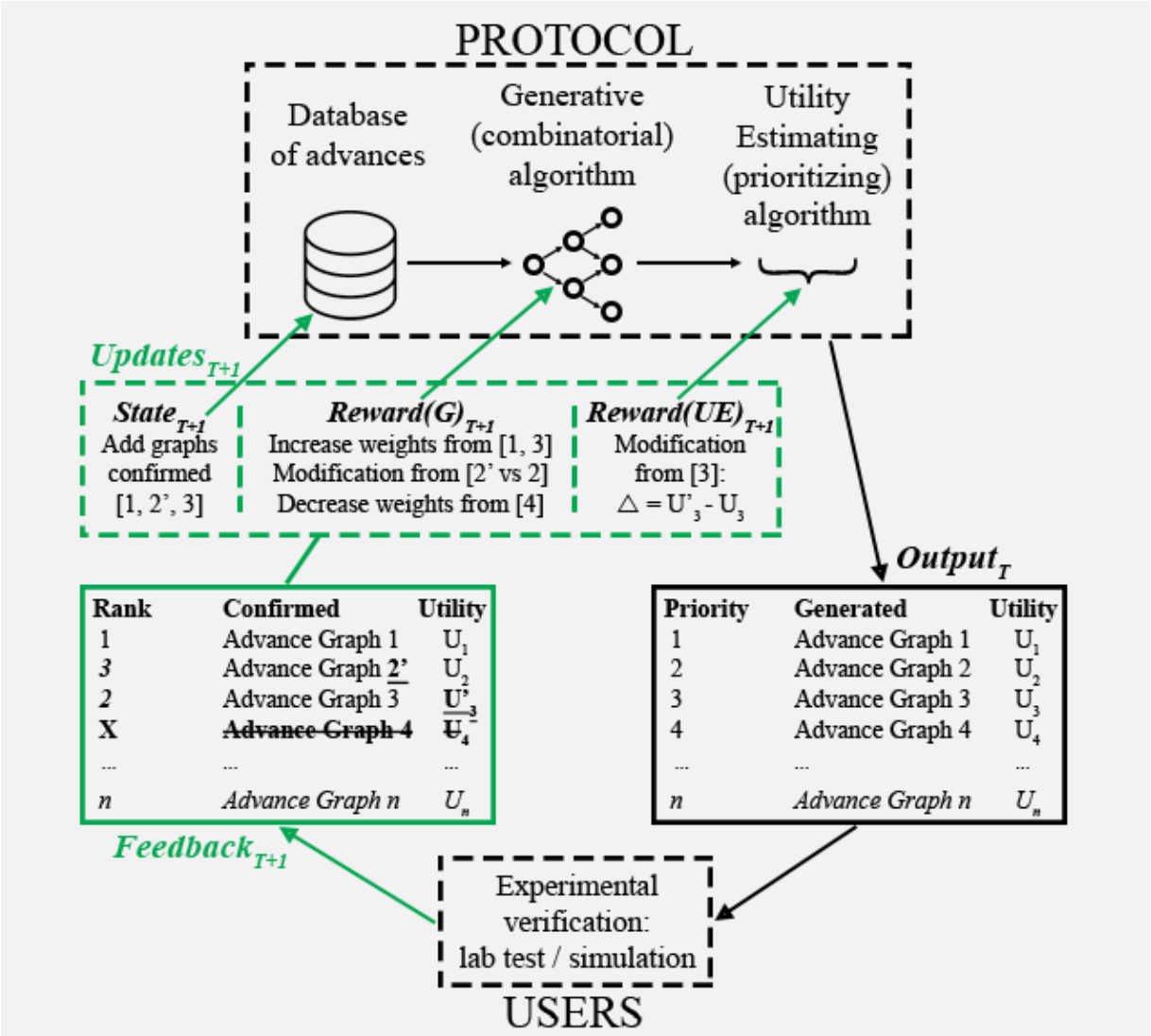


Figure 8. The reinforcement learning loop

Feedback at time T+1 is collected from experimental verification of the protocol's Output at time T, leading to the following Updates of the protocol database and algorithms at time T+1:
State$_{T+1}$: Valid (confirmed [1,3] and modified [2'] graphs) are added to the graph database
Reward(G)$_{T+1}$: The Generative algorithms that led to confirmed [1,3] or disproved [4] graphs receive increased or decreased weights, those that led to graph modifications [2'] are reevaluated according to the type and quantity of nodes and vertices that were modified
Reward(UE)$_{T+1}$: The Utility Estimating algorithms that led to modified utility [3] are reevaluated according to the loss (or gain) of utility [U'_3- U_3] and rank at time T+1 versus priority at time T.

These algorithmic improvements to maximize the expected rewards in the next iteration of the protocol must be programmed as an automated, self-defining policy to select the right architectures or discard the wrong algorithms. Integrated in the supervising Bayesian or evolutionary model-selecting algorithm, introduced in section 2.2, this policy would orient the selection of one or another of these algorithms, or modify them, depending on the rewards retrieved from experimental verification or invalidation of generated graphs and the actual versus estimated utility. In some cases, certain more challenging experimental results could trigger the intervention of the protocol's host team, when profound modifications of the algorithms architecture have to be implemented manually.

Through this **iteration** between the generation of advances from the algorithms and the feedback provided from researchers and engineers' experimental tests, the generating and utility-estimating algorithms should become more and more effective over time. Not only would the database of advances grow through experimental validation of generated graphs, but this RL virtuous circle would also improve accuracy in the generation of new potential technological and scientific theories.

With an ever-growing database of technologies to combine, and self-improving combining-prioritizing algorithms, it would seem that the more users the protocol aggregates the faster and faster it could lead us towards exponential scientific and technological progress. Section 3.2 further details how to attract more users in the protocol's community by remunerating the contribution of experimental feedback from R&D labs, scientists and researchers via a supporting system that could rely on the Blockchain technology.

3.2. Rewarding feedback providers to attract new users

As the users need to contribute their experimental feedback to update the protocol's database and algorithms through the RL system previously introduced, a financial compensation equivalent to their contribution should then be awarded back to the feedback providers. This retribution would serve a double purpose: rewarding the collaboration of current users to improve the protocol and create an incentive to attract new users.

The exact way to securely collect feedback from users and preserve the rights of scientists, engineers R&D laboratories obviously still need to be extensively defined, but feedback collection will assuredly be the one key to improve the effectiveness of the algorithms in the long run and ultimately, make the protocol's success. If the protocol can be implemented and the collaboration from multiple users can be secured, we could also imagine that a new form of intellectual property rights may be required to acknowledge the contribution of the protocol's development team, individual users, the broader protocol community, and the general public… but that is beyond the scope of this paper.

Aside from updating the core algorithms and growing the source graph database, feedback collection will also be key to compare, re-evaluate and consolidate advances graphs and utilities from various scientific and technological domains into a single system. Though individual users may keep private databases for their own use, it seems that the protocol's host team should maintain a well-updated, general, open and public graph database for the benefit of the larger community if the Singularity is to be effectively targeted. This consolidation activity should make the protocol's algorithms more comprehensive, aiming at the integration of **transfer learning** from a particular domain and situation that could come to be useful in a similar situation in another domain.

Just like feedback should be collected from experimental tests, a secondary support system (that may have to be handled manually) could be designed to reward users who provide other information, such as sets of advances that are not yet included in the general open database, or advances that are defined with more precise graphs and/or more precise utility vectors. Recouping multiple databases by crossing them over thanks to certain pivotal advances should not be extremely task in a later stage, when said advances or private databases can be retrieved, and properly rewarded.

Key core updates

- Add experimentally verified advance graph with features, relations and utility to the general database
- Increase/decrease weight of algorithm(s) that led to verified/disproved advance graph
- Alter algorithm(s) that led to successful but modified advance graph
- Alter algorithm(s) that led to discrepancy between estimated and actual utility

Key support updates

- Add new advance graph to the general database, from sources external to the protocol's generation/verification
- Consolidate similar advances with more or less precise graphs from multiple users and/or domains in the general database
- Consolidate utilities of similar advances from multiple users in the general database

Other support updates

- Consolidate simplified / extended graphs versions management: possibility to focus graphs on a user's specific area of concern and reduce computing power required
- Open confirmed algorithms to the general public

Figure 9. Non-exhaustive list of feedback triggering updates and cryptocurrency rewards

Since encryption and financial compensation will be key in this exchange of data, the Blockchain technology would seem to be a particularly well-fitting candidate to support feedback retrieval from multiple users and retribution with a cryptocurrency. By either relying upon the Ethereum Blockchain or creating a new Blockchain expressly designed to channel feedback to update the algorithms and database, the critical aspects of **anonymity** and **trust** would be ensured. Smart contracts should also conveniently provide a straightforward and integrated means to support the exchange of feedback against a compensatory cryptocurrency reward.

However, due to the large volume of data to be dealt with, it seems that the hosting of the database of advances graphs could not be handled directly through this Blockchain, as this technology currently seems unfit to handle large amounts of data transfers. Thankfully some decentralized data repositories like IPFS are already being developed to handle secure data repositories with inputs and outputs from multiple sources that can be conducted via a Blockchain. If the Blockchain can be used as a time seal of data transfer to the database host, the assurance of speed, security and access management to the database will be the only remaining aspects to ensure when selecting or developing the right solution for the general open graph database.

Any R&D lab, researcher and engineer that provides some feedback for the algorithms, database or other supporting activities of the protocol should be awarded an amount of the underlying cryptocurrency corresponding to their contribution. Though the exact amounts rewarded for each contribution still need to be defined in details, they should primarily be calculated from the reward signal users' experiments provide to the RL system, with the ultimate aim of rewarding the benefits shared to the whole community of users. These cryptocurrency rewards should also vary with the complexity and costs required to prove or disprove specific advances, or their importance to

consolidate or untangle the protocol's algorithms or database. Proving that a new concept of bicycle works is not the same challenge as proving the existence of the Higgs Boson

3.3. Securing initial development and long-term use through a financial support system

The present protocol breaks down the R&D process in a series of steps to be handled by existing AI algorithms. Due to the complexity of data and processes at hand, a significant amount of resources (time, expertise, computing power…) will be needed to be invested at first to develop the protocol into a minimum viable application. Even though some successful research projects already make use of a generative model previously presented[24] and could serve as a base to build the protocol upon, getting it up and running for a wide range of scientific and technological domains remains a very complex task. And this does not even include the additional challenges of securing feedback collection and integration for the long term and ensuring flexible mechanisms for evolution.

In this initial development phase, the protocol will need to be centralized and supervised, while the goal in the long term would be to make it as decentralized and unsupervised as possible. Though handling the protocol through a single (private) system may be a better solution, it seems that opening a general graph database and a library of algorithms to the general public could prove more beneficial for all to profit from individual users' resources and decentralize the protocol's costs as well as its benefits while consolidating its success through a technology like the Blockchain.

This question of efficiency and philosophy may be answered with different arguments for the long run, but in the short term, the initial development will be critical to ensure the protocol can be properly set up for the future. In any case, this initial phase will require significant funds (which can be obtained from various sources such as financial participations of beneficiaries of the protocol, traditional venture capital, an Initial Coin Offering - ICO - of the protocol's underlying cryptocurrency, etc.) that must be conceived as a viable form of investment in the long run for the stakeholders involved.

If the protocol's feedback collection can be properly handled by the Blockchain technology, as explained in section 3.2, several benefits from this technology could secure the protocol's success and financial soundness in the long run. In such a system, **the protocol's users, R&D labs, researchers, scientists and engineers**, providing feedback for the protocol's algorithms RL system would be assimilated to "miners" of the Bitcoin Blockchain. As their direct contribution to the improvement of the protocol's algorithms and technology database is compensated through cryptocurrency emission, users that provide feedback would be the backbone of the underlying cryptocurrency money supply in the long run.

As a side benefit, the use of the protocol and the retribution from its cryptocurrency would also help, in the long run, to alleviate the problem of the funding of fundamental science without direct commercial applications, which is especially critical for publicly funded R&D labs and researchers.

In addition to benefitting from the protocol's generation of viable potential advances, R&D labs would also be among the pillars and key beneficiaries of the financial ecosystem supporting the protocol, alongside its development and management stakeholders. As the two key groups involved the money supply, they could also work more closely together. Since certain experiments are more costly than others, and that certain fundamental bottlenecks in the protocol may require extra incentives to stimulate research to unlock them, other mechanisms could also be conceived through the collaboration of these two groups to increase or decrease the amount of cryptocurrency awarded for experimental tests or simulation.

In the long run, the protocol's **development and management stakeholders**, advances graphs database(s) host(s) and other parties involved in the consolidation of the database from different domains and technologies sub-sets, should also receive proper compensations of the underlying cryptocurrency. In the short term, the ICO (or "Reverse" ICO when the initial set up will be complete since ICOs have

recently been less effective source of funds) could be a convenient means to raise capital to initiate the works to launch the protocol and secure long-term return to early investors.

Whether the money supply should be capped, linear, exponential or logarithmic in the long run still needs to be properly defined, as well as the exact distribution scheme. Several methods could be conceived such as weighing or pegging the money supply with the world's economic growth, the growth of countries and/or companies where it is mostly used, technological progress, the impact the protocol may have, etc.

It would seems imperative, however, to include mechanisms to anticipate and limit daily volatility in the cryptocurrency's founding principles so as to prevent it to become impractical for day-to-day use. Some other systems may also be added to make the cryptocurrency flexible enough to evolve according to future conditions of the protocol, conditions of its use and users, or other economic situations which cannot be clearly anticipated.

In any case, the initial attribution mechanisms of the cryptocurrency during its launching phase and for its ICO should clearly be skewed to ensure developers, early contributors, investors and early supporters are properly rewarded in the long run. This could either happen by anticipating the cryptocurrency's appreciation or by integrating mechanisms of cumulative retribution for early stakeholders to benefit as the protocol and its underlying cryptocurrency gain prominence in the long run.

Discussion

Challenges

The Singularity Protocol introduces an organization and structure of concepts, algorithms and techniques that aim at reaching exponential scientific and technological progress with existing technologies and processes. Though it introduces a credible architecture and proposes solid technical candidates to achieve this goal, each and every step of the protocol still has to undergo thorough experimenting. Having yet to be validated through a battery of diverse tests before the protocol can yield any outcome, it is obvious that the actual implementation of the protocol beyond numerous technical hurdles could, and most probably will, reveal multiple problems and require multiple adjustments.

Some of these problems may stem from the general approach of grouping scientific and technological discoveries, inventions and innovations into the single concept of "advances". Others could arise from the necessary, and potentially problematic, articulation of the multiple disparate elements introduced to handle each task. More issues could also come from a possibly excessively ambitious generalization of simple concepts and algorithms or through the inherent challenges of graph networks, which are recapitulated in the last section of the GN paper.

Another major source of challenges will certainly arise from the complexity of numerous advances as their corresponding knowledge graphs could regroup hundreds, thousands or more nodes and relations. The computing power required to process such complex objects through the multiple processes introduces may be a critical technical limitation that could prevent any actual implementation of the protocol. Working with a large (and growing) database of such graphs, the protocol could just be structurally too complex to be expanded beyond very limited domains. The "brute force" approach of generating advances graphs in vast quantities could just go beyond the resources of any potential user, no matter how resourceful, at least for now.

Furthermore, due to the length of the current R&D process, another challenge may come from the time required between the generation of advances and the experimental verification. The weeks, months or even years that could be required to verify enough of the first advances generated by the protocol's

algorithms could disrupt the effectiveness of the RL process or render it impractical. This large sample efficiency issue will also certainly make initial investors impatient, if not doubtful, of any success in the long run, may lead to more non-technical sources of problems.

Other types of challenges could come from the actual advances that are being targeted, and more specifically the scientific and technological domains of concern. In certain domains upon which the protocol itself relies, which particularly include the domains of mathematics and computer programming, the protocol's algorithms could encounter some issues during the generation and/or the utility estimation. Since the protocol itself is tributary of certain algorithmic paradigms, it could become ineffective or less effective when these paradigms are more directly involved.

As a last potential issue among many other challenges that cannot be anticipated, the actual implementation of the protocol needs to be underlined. Though the ideal implementation of the protocol is sketched in this paper between an initial development team and a long-term decentralized collaboration, the actual implementation in the early phase, the transition to a decentralized application and the sustainability of the long-term use of the protocol all may create troubles that have not been anticipated. Since the devil is in the detail, many new challenges will undoubtedly appear as the protocol is made a reality, and only a most dedicated and resolved team, ready to endure many unexpected difficulties in the long run can eventually succeed.

Open questions

Some final considerations are more philosophical in nature. A first one comes from the very unlikely, yet possible, limit of automated combinatorial creativity. Combining existing advances, features and relations could lead to dead end after the generation of only a limited number of new useful advances. Even though the very idea of an entirely new concept is also subject to questioning, if entirely new advances are not periodically added to the database the protocol may be of little use with an overwhelming number of useless graphs.

Another unlikely scenario could come in the long term use of the protocol. Though the protocol targets exponential growth of science and technology, it could also happen to lead towards a more or less distant limit, exhibiting an S-shaped curve rather than an exponential one. Again, if new advances, features and relations are not periodically added to the original pool upon which the protocol's algorithms draw to generate advances, a limit could be reached in its generative capacity.

In a more positive note, another consideration comes from the potential success of the protocol. If it does lead toward new technologies, it could be conceivable that the protocol's approach would be adaptable beyond science and technology. Some other fields where a similar approach could be applicable may include architecture, design, urban planning or even art. And, even though the protocol is designed to reach the Technological Singularity without going through the creation of an artificial general intelligence, its success could also be a stepping stone towards the creation of the AGI.

Conclusion

Even though it presents some important hurdles and limitations that will need to be judiciously tackled, the Singularity Protocol proposes a conceptual and speculative, but also structured and credible approach to generating new scientific and technological theories at scale, through artificial intelligence. Though ambitious in its goal of reaching exponential technological progress, this protocol breaks down the theoretical process of R&D into a sequence of simple steps, each matched with solid algorithmic and technological candidates, so as to be handled by machines.

An effort has been made to describe each section in the simplest terms with the aim of facilitating comprehension for a wider audience with minimal engineering background, while providing sources

and references to the complete underlying scientific, mathematical and algorithmic material for the more advanced readers.

The objective of this paper is to spark interest and discussions on the Technological Singularity, which already seems to be within reach, if not to actually launch the protocol's deployment, provided the required human, financial and computing resources can be harnessed. Do not hesitate to contact the author if you are interested in collaborating towards this goal.

References

[1] Albert Einstein (1954) "Ideas and opinions" – see letter from 1945 in response to the publishing in the same year of "An Essay on the Psychology of Invention in the Mathematical Field" by Jacques S. Hadamard

[2] Peter W. Battaglia, Jessica B. Hamrick, Victor Bapst, Alvaro Sanchez-Gonzalez, Vinicius Zambaldi, Mateusz Malinowski, Andrea Tacchetti, David Raposo, Adam Santoro, Ryan Faulkner, Caglar Gulcehre, Francis Song, Andrew Ballard, Justin Gilmer, George Dahl, Ashish Vaswani, Kelsey Allen, Charles Nash, Victoria Langston, Chris Dyer, Nicolas Heess, Daan Wierstra, Pushmeet Kohli, Matt Botvinick, Oriol Vinyals, Yujia Li, Razvan Pascanu (2018) "Relational inductive biases, deep learning, and graph networks" *arXiv:1806.01261*

[3] See previously mentioned paper in note ([2]), page 21, and the references cited in the paper

[4] See previously mentioned paper in note ([2]), page 21, and the references cited in the paper

[5] See previously mentioned paper in note ([2]), page 9, and the references cited in the paper

[6] See previously mentioned paper in note ([2]), page 11

[7] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, Jeffrey Dean (2013) "Distributed Representations of Words and Phrases and their Compositionality" *arXiv:1310.4546* - See also TensorFlow Tutorial: Vector Representations of Words

[8] Hanjun Dai, Bo Dai, Le Song (2016) "Discriminative Embeddings of Latent Variable Models for Structured Data" *arXiv:1603.05629*

[9] Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, George E. Dahl (2017) "Neural Message Passing for Quantum Chemistry" *arXiv:1704.01212*

[10] Wengong Jin, Regina Barzilay, Tommi Jaakkola (2018) "Junction Tree Variational Autoencoder for Molecular Graph Generation" *arXiv:1802.04364*

[11] See previously mentioned paper in note ([2]), page 16, and the references cited in the paper

[12] See previously mentioned paper in note ([2]), page 17-19, and the references cited in the paper

[13] Yujia Li, Oriol Vinyals, Chris Dyer, Razvan Pascanu, Peter Battaglia (2018) "Learning Deep Generative Models of Graphs" *arXiv:1803.03324*

[14] See previously mentioned paper in note ([13]), page 3

[15] See previously mentioned paper in note ([10]), page 7

[16] Aleksandar Bojchevski, Oleksandr Shchur, Daniel Zügner, Stephan Günnemann (2018) "NetGAN: Generating Graphs via Random Walks" *arXiv:1803.00816*

[17] Jianmin Bao, Dong Chen, Fang Wen, Houqiang Li, Gang Hua (2017) "CVAE-GAN: Fine-Grained Image Generation through Asymmetric Training" *arXiv:1703.10155*

[18] See previously mentioned paper in note ([2]), page 16, and the references cited in the paper

[19] Hanxiao Liu, Karen Simonyan, Oriol Vinyals, Chrisantha Fernando, Koray Kavukcuoglu (2017) "Hierarchical Representations for Efficient Architecture Search" *arXiv:1711.00436*

[20] See previously mentioned paper in note ([2]), page 10

[21] See previously mentioned paper in note ([2]), page 16

[22] See previously mentioned paper in note ([2]), page 16 – Refer to section 3.2 for details on the notations and section 4 for details about the two extracted variants and the other GN Blocks architectures.

[23] See previously mentioned paper in note ([2]), page 19

[24] See previously mentioned paper in note ([10])