A Knowledge Representation of the Miller-Urey Experiment

Javier Burgos[1,2]

[1] Research Department, Corporación para la investigacion y la innovación-CIINAS, La Calera, Cundinamarca, Colombia
[2] Centro de investigacion en informática del derecho-CIDI, Universidad Externado de Colombia, Bogotá, Colombia

E-mail: jdburgoss@corporacionciinas.org

Abstract

We present an ontology log-OLOG- representation of the classical Miller-Urey experiment, usually considered as a paradigm for spontaneous generation of biomolecules on the prebiotic Earth and also, as a key in understanding the chemical evolution phenomena linked to the origins of life. Ologging The Miler-Urey experiment enables us, through the categorical notion of fibre product or pullback, to define the concept of Biogenic Space, as a space containing low complexity biogenic units subjected to appropriate physical and chemical conditions, facilitating the synthesis of highly complex organic molecules. Also, we characterize the Biogenic Space as a concrete universal object that could be associated with the preconditions for life inside various structures in the universe such as exoplanets and exomoons located in habitable zones, but also in interstellar and intergalactic organic clouds.

Keywords: Preconditions of life, Miller Urey experiment, Knowledge representation, Ontology log, Universality.

1. introduction

In 1959, Stanley Miller and Harold Urey [1] reported on the spontaneous formation of glycine when applying an electric discharge on a mixture of simple molecules, giving birth to modern research on the origins of life. Since then, a number of different hypotheses have been formulated to explain the synthesis, from simple molecules (water, ammonia, methane, carbon oxides), of simple organic molecules (formaldehyde, hydrogen cyanide, formic acid), and from the latter ones to biological monomers (amino acids, purines, pyrimidines) up the ladder of the complexity of life. An interesting aspect of the experiment was that the general relative abundance of amino acids was remarkably similar to which is found in certain carbonaceous meteorites.

It is important to note that, from the analysis of organic matter in meteorites and other solar system objects, we now know that abiotic synthesis can create a wide range of organic compounds far beyond those found on Earth. This hints that the degrees of complexity and diversity of prebiotic organics are much larger than previously thought [2].

In the Miller–Urey experiment, plasma from the electric discharge creates ions as well as radicals (molecules or atoms carrying an unbound, single electron). As the plasma temperature drops the molecules in the discharge combine to form new substances. In the process radicals that combine with other radicals will tend to form new, non-radical species. However, radicals reacting with non-radicals will lead to new radicals that may react further. While small molecules with elevated vapor pressure essentially remain in the gas phase and remain subject to further discharge, larger and more polar molecules are more likely to dissolve in the liquid phase where they may react further.

Despite having a large number of variations of the Miller-Urey experiment [3], they all have a series of common characteristics that allow them to be classified within the same framework, those related with the experimental basis of the theory of the evolution of chemistry in the universe, a key to understanding the origin of life [4].

Recently, Kauffman and collaborators [5] present a theory for the evolution of molecule masses and show that small molecules grow by random diffusion and large molecules by a preferential attachment process leading eventually to life's molecules. It reproduces correctly the distribution of molecules found via mass spectroscopy for the Murchison meteorite and estimates the start of chemical evolution back to 12.8 billion years following the birth of stars and supernovae. Also, using the scaling of reaction rates with the distance of the molecules in space, they recover correctly the few days emergence time of amino acids in the Miller-Urey experiment.

Moreover, Wollrab and Ott [6] were able to demonstrated that a miller–Urey broth mirrors the mass density distribution of all Beilstein indexed organic molecules, leading to the conclusion that the Miller–Urey experiment seems to explore the entire range of prominent, stable masses of molecules that are part of the Beilstein/Reaxys database and made from the same atomic elements.

These results suggest that the Miller-Urey experiment could have a universal character, which is what is explored in this article using a knowledge representation approach since the mathematical category theory.

2. A Categorical Representation of the Miller-Urey Experiment

In this paper, we will use the "ontology log" or olog as a possibility for such a framework [7]. Ontology is the study of what something is, i.e. the nature of a given subject, and ologs are designed to record the results of such a study. The structure of ologs is based on a branch of mathematics called category theory. An olog is roughly a category that models a given real-world situation.

Following Spivak [8,9], a basic olog is a category in which the objects and arrows have been labelled by English-language phrases that indicate their intended meaning. The objects represent types of things, the arrows represent functional relationships (also known as aspects, attributes, or observables), and the commutative diagrams represent facts.

On the other hand, Bottcher [10] defined biogenic units as the ordered pair (Cm, Ci), where Cm represents the molecular complexity, and Ci represent the informational complexity. Since this point of view, biogenic units can correspond to individual molecules, multimeric complexes, protocells such as LUCA, organisms and even societies. For the present work, regarding the Miller Urey experiment, we define simple molecules like water, ammonia, methane, carbon oxides as low complexity biogenic units, and the final products of the experiment like amino acids, purines, pyrimidines among others as high complex biogenic units. It is clear that the

adopted definition of biogenic units applied to the complex organic molecules with ubiquitous presence in the well-known Murchison, Murray and Orgueil meteorites [11].

2.1 Ologging Miller-Urey experiment

In order to represent the Miller Urey experiment, three objects are defined, which are an integral part of it, first, C and N based simple molecules like water, ammonia and methane among others. Second, the appropriate physical conditions like electric discharges and elevated vapour pressure. Third, in any Miller-Urey type experiment, as final products, some biological monomers like amino acids and nitrogenous bases are obtained. The causal relationship between these objects, here represented as boxes, is depicted by arrows representing the observables or aspects of the Miller Urey experiment (Figure 1). Thus, the diagram or graph for the olog of the experiment is the following:
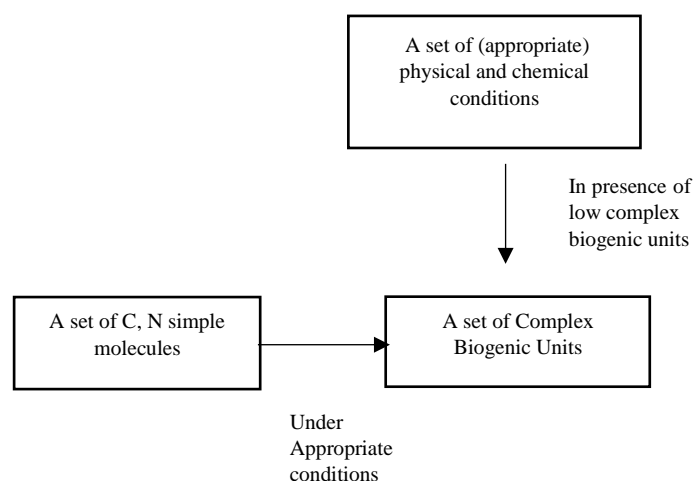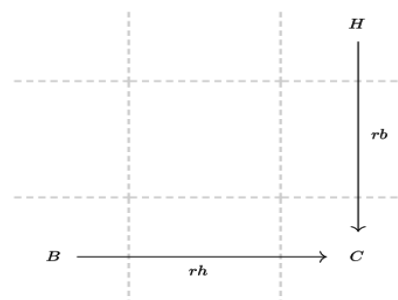

Figure 1. The Miller Urey olog.

This olog represent, in a simple manner, a typical Miller Urey experiment and can be expressed as short diagram as follows:

Where H represent a set of physical and chemical conditions, B represent the set of low complexity biogenic units, r_b and r_h correspond to the acting functions producing C, the set of complex biogenic units.

2.2 The Miller-Urey Pullback: the notion of Biogenic Space

The Miller-Urey olog represent, in a simple manner, a typical Miller-Urey experiment but, in addition, following the formalism of the category theory, it allows defining a new concept, that of Biogenic Space, through the use of the fibre product or pullback, as presented in the following diagram (figure 2):
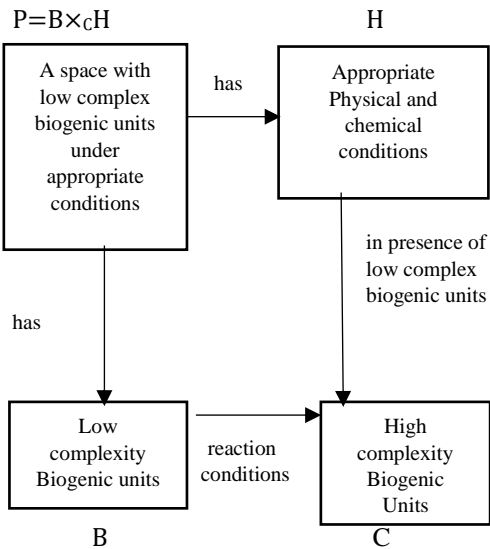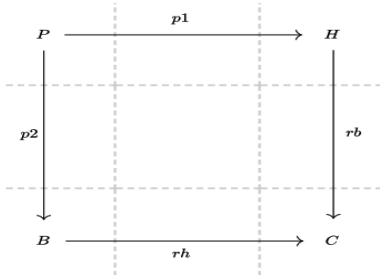
$P=B\times_C H$

Figure 2. Olog of the Biogenic Space as a fibre product

In accordance with, a space (P) with low complexity biogenic units subjected to appropriate physical and chemical conditions, facilitating the synthesis of highly complex organic molecules is a Biogenic Space.

Technically, a pullback (also called a fibre product, fibre product, fibered product or Cartesian square) is the limit of a diagram consisting of two morphisms $r_h: B \to C$ and $r_b: H \to C$ with a common codomain. The pullback is often written $P=B\times_C H$ and can be represented by a concise diagram as follows:
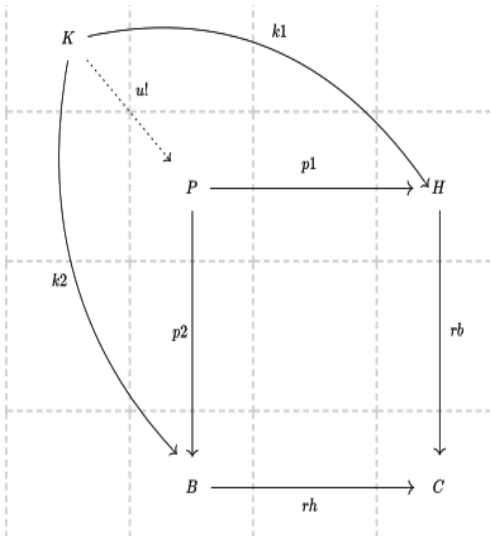

$P=B\times_C H$ comes equipped with two natural morphisms or projections $p_1:P \to H$ and $p_2:P \to B$.

3. The Universality of Biogenic Fibre Products

Lemma (Universal property of Biogenic Spaces): Suppose given the diagram of objects and morphisms of a Miller-Urey experiment as below:
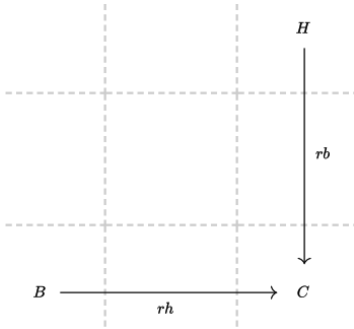
For any set K and commutative solid arrow diagram as follows:

There exists a unique arrow u!: K→ P, (P=B×$_C$H) making everything commute i.e. $k_1= p_1°u$; $k_2=p_2°u$. these compositions seems indicate that to obtain a pool of complex biogenic units as a result of a Miller-Urey experiment there can be two ways, the first having previously the appropriate physicochemical conditions and then adding the low complexity biogenic units and the second, which consists of having initially the low complexity molecular species which later become exposed to the appropriate reaction conditions.

Moreover, the Biogenic Space, as with all universal constructions [8], if it exists, is unique up to isomorphism. In fact, given two pullbacks (X, x_1, x_2) and (Y, y_1, y_2) of the same cospan $B \to C \leftarrow H$, there is a unique isomorphism between X and Y respecting the pullback structure.

4. Discussion

The present work is oriented towards the study of the preconditions for life, which can become as overwhelming as the study of the appearance and development of life itself, the latter being the main object of astrobiology and the former the main object of astrochemistry. Recently, Kauffman, Jelenfy and Vattay [5], proposed a theory about the evolution of the molecular composition of the universe, taking as a basis the concentrations of complex organic molecules resulting from a Miller-Urey type experiment, from which they managed to simulate the distributions of organic molecules found in the immense interstellar and intergalactic clouds.

On the other hand, Wollrab and Ott [6] describe how a broth resulting from a Miller-Urey experiment reflects the mass-density distribution of all organic molecules in the Beilstein index, a result which suggests, along with that of Kauffman and collaborators, that the Miller-Urey experiment has a universal character. Moreover, given that the gases needed for effective prebiotic syntheses are wide spread in the Universe, and the robust nature of the Miller-type synthetic reactions, simple prebiotic compounds are likely common in the cosmos. It is likely that the formation of more complex molecules, including one with rudimentary capabilities of imperfect replication, could also be common throughout the Universe [12].

Inspired by these results, we suggest, using a knowledge representation strategy, specifically ontology logs or Ologs from category theory, that Miller-Urey type experiments contain spaces, that we have called biogenic spaces, which have the character of concrete universals, that is, spaces whose macroscopic behavioural properties, regardless of scale, are maintained even when their specific microscopic components can be replaced by others. From this perspective, biogenic spaces can be circumscribed to the flask of a laboratory experiment [1], the primitive earth [13], the moon Enceladus [15], the habitable zone around a star [14], the galactic habitable zone or the immense clouds of interstellar and extra galactic organic molecules [16], which, if considered as topological spaces or smooth manifolds, can be studied using the powerful mathematical formalism that is an integral part of current astrophysics and cosmology. But also, important, Biogenic Spaces can be objects or observables, common in the universe, which allows us to think that their ubiquity is a precondition, at least necessary, for the subsequent appearance of life as a phenomenon (apparently) singular.

Acknowledgements

References

[1] Miller S and Urey H 1959 *Science* **130** 245

[2] Meringer M Cleaves H 2017 *Phil.Trans.R.Soc.* **A375** 20160344

[3] Parker E Cleaves J Burton A Glavin D Dworkin J Zhou M Bada J Fernández F 2014 *J. Vis. Exp*. **83** e51039

[4] Coveney P Swalding J Wattis J and Greenwell H 2012 *Chem. Soc. Rev* **41** 5430

[5] Kauffman S Jelenfy D and Vattay G 2020 *J theoretical Biol* **486** 110097

[6] Wollarb E and Ott A 2018 *New J. Phys*. **20** 105003

[7] Spivak D Kent R 2012 *PLoS ONE* **7** e24274.

[8] Spivak D Category Theory for Scientists 2014 The MIT Press

[9] Maclane S 1971 Categories for the Working Mathematician springer

[10] Bottcher T 2017 *J. Mol. Evolution* **86** 1

[11] Biver N and Bockelée-Morvan D 2015 *Proceedings of the International Astronomical Union* **11** 228

[12] Bada J 2013 *Chem. Soc. Rev*. **42** 2186

[13] Miller S 1953 *Science* **117** 528

[14] Kahana A Schmitt-Kopplin P Lancet D 2019 *Astrobiology* **19**

[15] Cockell C Bush T Bryce C et.al. 2016 *Astrobiology* **16** 1

[16] Tielens A 2013 *Review of Modern Physics* **85** 1021

Appendix

Briefly [9], a category C can be described as a set Ob, whose members are the objects of C, satisfying the following three conditions:

Morphism: for every pair X, Y of objects, there is a set Hom (X, Y), called the morphisms form X to Y in C. If f is a morphism from X to Y, we write $f{:}X{\rightarrow}Y$.

Identity: for every object X, there exist a morphism id_X in Hom (X, X), called the identity on X.

Composition: For every triple X, Y and Z of objects, there exists a partial binary operation from Hom (X, Y) $\times$ Hom (Y, Z) to Hom (X, Z), called the composition of morphisms in C. If $f{:}X{\rightarrow}Y$ and $g{:}\ Y{\rightarrow}Z$, the composition of f and g is notated $(g{\circ}f){:}X{\rightarrow}Z$.
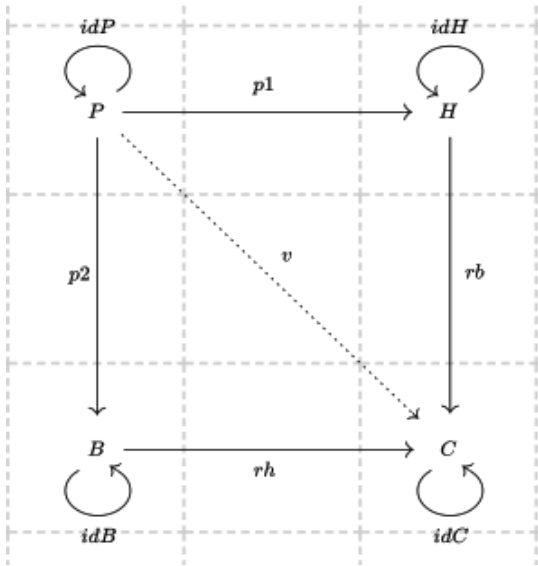
Identity, morphisms, and composition satisfy two axioms:

Associativity:
If $f{:}X{\rightarrow}Y$, $g{:}\ Y{\rightarrow}Z$, and $h{:}\ Z{\rightarrow}W$, then $h{\circ}(g{\circ}f)=(h{\circ}g){\circ}f$

Identity: If $f{:}X{\rightarrow}Y$, then $(id_Y{\circ}f)=f$ and $(f{\circ}Id_X)=f$.

Consider the diagram below named MU, whose objects are P, H, B, C and the morphisms are as depicted in the diagram, it is easy to verify that MU is a category.

Having established MU as a category, we will proceed to demonstrate the Lemma of the Universal property of Biogenic Spaces. The central element of the proof is to see that u:K in P is unique. Let us suppose, on the contrary, that there are u_1 and u_2 morphisms of K in P. Now, as p_1 is the projection of the component of specific and carefully controled physicochemical conditions in the biogenic space, and p_2 the projection of the component of simple molecular compounds, present in specific concentrations in the experiment, it must be then that $k_1 = p_1 o u_1 = p_1 o u_2 = k_1$, and also that $k_2 = p_2 o u_1 = p_2 o u_2 = k_2$, which implies that $u_1 = u_2$, which contradicts the initial hypothesis.