Using Field Equations To Determine The Meaning Of Informal Texts

Philippe Larvet
Alcatel-Lucent, Bell Labs France, Centre de Villarceaux, Route de Villejust, 91620 Nozay, France
Email: Philippe.Larvet@alcatel-lucent.fr

Abstract
Einstein Field Equations (EFE) are used within the Theory of General Relativity to determine the curvature of space-time resulting from the presence of mass and energy. That is, EFE determine the metric tensor of space-time for a given arrangement of stress-energy in the space-time. An informal text, containing phrases and words, can be seen as a portion of space where words determine a *semantic curvature* resulting from their presence and their meaning. That is, EFE can be used to determine a sort of *semantic tensor* or *semantic field* of text-space for a given arrangement of word-semantics within the area of an informal text. This paper describes an original use of Field Equations in order to compute the meaning of informal texts, leading to the ability to determine and compare the significance of any kind of informal documents (articles, book chapters, mails, technical documentation, etc.), even if written in different languages.

Keywords
Field equations, semantic tensor, semantic field, informal text, text analysis, natural language processing

1. EINSTEIN FIELD EQUATIONS

Einstein field equations (EFE) or Einstein's equations are a set of ten equations in Einstein's Theory of General Relativity in which the fundamental force of gravitation is described as a curved spacetime caused by matter and energy [1] They were first published in 1915 [2]. These dynamic equations describe how matter and energy modify the space-time geometry. The correlated geometry curvature around a matter source is interpreted as the gravitational field of this source.

The EFE collectively form a tensor equation and equate the curvature of spacetime (as expressed using the Einstein tensor) with the energy and momentum within the spacetime (as expressed using the stress-energy tensor).

The EFE are used to determine the curvature of spacetime resulting from the presence of mass and energy. That is, they determine the metric tensor of spacetime for a given arrangement of stress-energy in the spacetime. Because of the relationship between the metric tensor and the Einstein tensor, the EFE become a set of coupled, non-linear differential equations when used in this way.

The Einstein field equations (EFE) may be written in the form:[1]

$$R_{ab} - \frac{1}{2}R\,g_{ab} = \kappa T_{ab} = -\frac{8\pi G}{c^4}T_{ab}.$$

where R_{ab} is the Ricci tensor, R the scalar curvature, g_{ab} the metric tensor and T_{ab} the stress-energy tensor. The constant κ (kappa) is called the *Einstein constant (of gravitation)*, where π (pi) is Archimedes' constant, G the gravitational constant and c the speed of light.
The above form of the EFE is for the −+++ metric sign convention, which is commonly used in general relativity, and which is used by convention here.

The change of sign on the right hand side occurs because the values of T_{ab} have signs which are determined by the sign convention. On the other hand, the values of the left hand side are convention independent: R_{ab} has values which are independent of the convention because the convention dependencies of R and g_{ab} cancel out.

The EFE is a tensor equation relating a set of symmetric 4 x 4 tensors. It is written here using the abstract index notation. Each tensor has 10 independent components. Given the freedom of choice of the four space-time coordinates, the independent equations reduce to 6 in number.

Although the Einstein field equations were initially formulated in the context of a four-dimensional theory, the equations hold in n dimensions. The equations in contexts outside of general relativity are still referred to as the Einstein field equations (if the dimension is clear).

Despite the simple appearance of the equation it is, in fact, quite complicated. Given a specified distribution of matter and energy in the form of a stress-energy tensor, the EFE are understood to be equations for the metric tensor g_{ab}, as both the Ricci tensor and Ricci scalar depend on the metric in a complicated nonlinear manner. In fact, when fully written out, the EFE are a system of 10 coupled, nonlinear, hyperbolic-elliptic partial differential equations.

One can write the EFE in a more compact form by defining the Einstein tensor which is a symmetric second-rank tensor that is a function of the metric:

$$G_{ab} = R_{ab} - \frac{1}{2} R g_{ab},$$

This point will be developed below.

2. A NON-STANDARD USE OF EINSTEIN FIELD EQUATIONS.

The purpose of this paper is to propose the description of the universe of a text expressed in natural language (NL) with the help of a tensor field equation.

A text is made of sentences, each of one being made of nodes, i.e. words or locutions.
To each node is associated a quatuor Q, described by a set of coordinates z, r, θ, γ, where

- z is the value of the signification of the node, taken in a specific dictionary
- r and θ are 2D polar coordinates of the node in the space-plan of the text,
- γ is the value of the grammatical role of the node in the sentence.

A tensorial analysis of the text mainly corresponds to the analysis of nodes, in order to elaborate a metric for every phrase. This metric is obtained through the resolution of a tensor equation taking into account all the nodes of the sentence.

Within the context of natural language processing, the problem is to compare the meaning of two different texts. More precisely, the problem we want to solve is to determine (to calculate) the meaning of the two texts, and to compare these meanings.

The basic idea is to calculate the "significance field" (champ de signification) of a given text

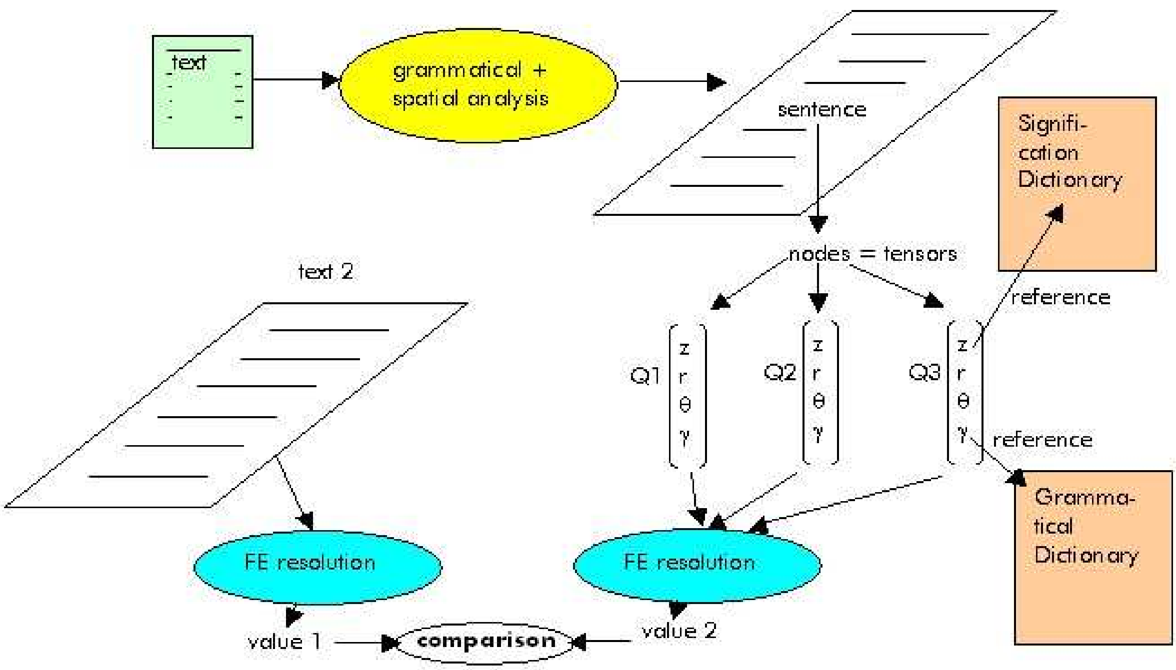
1) by considering the text as a short portion of a finite space, and

2) 2) by using a tensorial field equation to calculate the value of the "field of significance" (<<semantic field>>) for this text.

So, we come to an "evaluation function" that expresses a value for the meaning of the text (likewise a chess evaluation function evaluates a position in chess). Thus, by comparing the two values for two different texts, it is possible to compare the meanings of the two texts and then, to determine if they have neighbouring significations.

3. THE PROPOSED SOLUTION

The solution is based upon 3 basic axioms:
1) a text is considered as a finite portion of space;
2) the sentences composing the text are expressed with tensors [3].
3) the Einstein Field Equations used in physics are applied onto these tensors in order to determine the value of a "significance field" that represents the meaning of the text. The comparison of the values of these fields for two distinct texts allows to determine if these texts have neighbouring meanings.



The tensorial analysis of the text mainly consists in node analysis, in order to elaborate a "metric" for each sentence. This "metric" is obtained by the resolution of a tensorial equation using all the nodes of the sentence.

4. GENERAL RESOLUTION OF FIELD EQUATIONS

The mechanism of the resolution is based on the solutions to the standard Field Equations (FE) used in physics and cosmology (see for example http://io.uwinnipeg.ca/~vincent/4500.6-001/Cosmology/Field-Equations.htm). Within the original context, these FE describe how stress-energy causes curvature of space-time. Within the specific context of the meaning of a NL text, the FE is intended to describe how the "energy" of words causes the significance of the space represented by the text.

The FE are usually written in tensor form (using abstract index notation) as

$$G_{ab} = \kappa T_{ab}$$

where G_{ab} is called "Einstein tensor", T_{ab} is the stress-energy tensor and k is a constant whose the value is $8\pi G/c^4$, where G is the gravitational constant and c the speed of light – see [5] for example.

The tensor G_{ab} is related to the curvature of space-time and is a function only of the metric tensor and its first and second derivatives. The stress-energy tensor T_{ab} , which is the source of the gravitational field, includes stress (pressure and shear), density of momentum, and density of energy including the energy of mass (the source for Newtonian gravity).

It is possible to write the FE in a more compact way by redefining the tensor G_{ab} :

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}$$

which is, like T_{ab} , a rank-2 symmetric tensor, and is depending on the metrics. By working in a geometrical unit where G = c = 1, the relation is written:

$$G_{\mu\nu} = 8\pi T_{\mu\nu}$$

Within the original context of this FE, the left part of this equality represents the curvature of space-time such as it is determined by the metric, and the expression on the right represents the content mass/energy of space-time. Then, this equality can be interpreted as a set of equations describing how the curvature of space-time is linked to the content mass/energy of the universe.

Within the context of the analysis of a NL text, the equality expresses how the significance of the text is linked to the content of the nodes, in terms of meaning, grammatical role, and spatial position of nodes in the sentence.

5. FIELD EQUATIONS RESOLUTION WITHIN THE CONTEXT OF NATURAL LANGUAGE ANALYSIS

The tensorial analysis of a NL text consists in node analysis, in order to elaborate a metric for each sentence. This metric is obtained by the resolution of the tensorial equation using all the nodes of the sentence.

In a general way, the solutions of the FE are the "metric tensors" of space-time. They are often called "metrics" ("métriques"). They describe the structure of space-time by including the inertial movement of the objects, then by taking into account some specific space-time parameters like homogeneity, isotropy, expansion or contraction.

Considering the fact that the field equations are not linear, they cannot be completely solved, i.e. without doing some approximations (for example, there is no known complete solution for a space-time containing two massive elements, corresponding to the theoretical model of a twostars binary system for example). The study of the solutions to the FE is one of the activities of the cosmology.

Then, within the context of solving the FE for NL analysis, a specific metric has to be chosen, taking into account a "universe" (the text) with singular characteristics: homogeneity, complete isotropy, coefficient of expansion =0, coefficient of contraction =0. So, the metric of Friedmann-Lemaître-Robertson-Walker (FLRW, that is considered a specialization of Friedmann's equations – see [6], first writing of Friedmann's equations) is a good candidate, because it allows to describe a locally homogeneous, locally isotropic and non expansional universe.

With this metric, the solution of the FE can be written, in polar coordinates (r, θ, ϕ):

$$\boxed{\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a(t)^2\left(\frac{\mathrm{d}r^2}{1-kr^2} + r^2\mathrm{d}\Omega^2\right)}$$

where

- a(t) is the scale factor of the universe at the time t,
- k is the spatial curvature and can take 3 values = +1 (curve closed space corresponding to a spheric geometry), 0 (flat universe corresponding to the usual euclidian geometry) or –1 (curve open space corresponding to an hyperbolic geometry),
- $\mathrm{d}\Omega^2 = \mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2$ expresses the contributions of the metrics linked to the "direction" (θ,ϕ).

As we are within the context of a flat universe (the geometry of a text is neither spheric nor hyperbolic…), the metric can be written:

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a(t)^2(\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2)$$

So, the value of the local "significance field" $g_{\mu\nu}$ for a sentence of a text, described by the quatuor Q(z, r, θ, γ) is given by the matrix:

$$g_{\mu\nu} = \begin{pmatrix} -dt^2/z & 0 & 0 & -dt^2/dr^2 \\ 0 & a(t)^2 z dr^2 & 0 & 0 \\ 0 & 0 & a(t)^2 z r^2 d\Omega^2 & 0 \\ 0 & 0 & 0 & \gamma dr^2 \end{pmatrix}$$

This solution allows calculating the meaning of all sentences of the text, and the complete significance of the whole text.

6. References

1. Einstein, Albert (1916). "The Foundation of the General Theory of Relativity" (PDF). *Annalen der Physik*.

2. Einstein, Albert (November 25, 1915). "Die Feldgleichungun der Gravitation". *Sitzungsberichte der Preussischen Akademie der Wissenschaften zu Berlin*: 844-847. Retrieved on 2006-09-12.

3. We remind that, within the context of physics or linear algebra ("algèbre linéaire") a tensor ("tenseur") is a mathematical object that generalizes the notion of vector by allowing to embody the set of the representations, in all the 'bases' of a given vectoriel space (expace vectoriel) of a vector $\vec{u}$ described by its unitary components (u1, u2, u3) for example. A 'base', within the context of linear algebra ("algèbre linéaire") is a family of vectors that allow to be located in space through algebric coordinates.

4. Gamow, George (April 28, 1970). *My World Line : An Informal Autobiography*. Viking Adult. ISBN-10: 0670503762. Retrieved on 3-14-2007.

5. *Introduction to General Relativity* (Ronald Adler, Maurice Bazin, Menahem Schiffer), Mcgraw-Hill, June 1975 - ISBN 0070004234.

6. Alexandre Friedmann *Über die Krümmung des Raumes, Zeitschrift für Physik* 10, 377-386 (1922)