Acta Astronautica 78 (2012) 20–25

Contents lists available at ScienceDirect

Acta Astronautica

journal homepage: www.elsevier.com/locate/actaastro

The DISC Quotient

John R. Elliott [a,*], Stephen Baxter [b]

[a] Creative Technology Research Group, Leeds Metropolitan University, Leeds, LS6 3QS, England, UK
[b] c/o Ralph Vicinanza, 303 W 18th Street, New York, NY 1001, USA

ARTICLE INFO

Article history:
Received 1 March 2011
Accepted 5 March 2011
Available online 13 April 2011

Keywords:
Communication
Decipherment
DISC Quotient
Dissemination mechanisms and rationales
Dynamic analysis
Linguistic analysis
Post detection decipherment strategy
Signal analysis
Societal impact

ABSTRACT

D.I.S.C: Decipherment Impact of a Signal's Content.

The authors present a numerical method to characterise the significance of the receipt of a complex and potentially decipherable signal from extraterrestrial intelligence (ETI). The purpose of the scale is to facilitate the public communication of work on any such claimed signal, as such work proceeds, and to assist in its discussion and interpretation.

Building on a "position" paper rationale, this paper looks at the DISC quotient proposed and develops the algorithmic steps and comprising measures that form this post detection strategy for information dissemination, based on prior work on message detection, decipherment. As argued, we require a robust and incremental strategy, to disseminate timely, accurate and meaningful information, to the scientific community and the general public, in the event we receive an "alien" signal that displays decipherable information. This post-detection strategy is to serve as a stepwise algorithm for a logical approach to information extraction and a vehicle for sequential information dissemination, to manage societal impact.

The "DISC Quotient", which is based on signal analysis processing stages, includes factors based on the signal's data quantity, structure, affinity to known human languages, and likely decipherment times. Comparisons with human and other phenomena are included as a guide to assessing likely societal impact. It is submitted that the development, refinement and implementation of DISC as an integral strategy, during the complex processes involved in post detection and decipherment, is essential if we wish to minimize disruption and optimize dissemination.

1. Introduction

In this paper we present a strategy to follow the receipt of a complex and potentially decipherable signal from extraterrestrial intelligence (ETI). This includes a signal-processing algorithmic procedure that contains analysis stages based on the signal's data quantity, information-theoretic and linguistic structure, and affinity to known human languages. In addition, we propose a numerical scoring system based on the procedure's algorithmic steps to characterize the significance of the signal and its subsequent analysis for the purpose of public communication. This "DISC Quotient" scale is modeled on the example of "Rio Scale" for characterizing the discovery of an ETI.

2. DISC rationale

2.1. Suppose SETI succeeds

It is an indisputable fact that positive identification of a signal from an extraterrestrial source will have a

* Corresponding author. Tel.: +44 113 812 7379.
E-mail addresses: J.Elliott@leedsmet.ac.uk (J.R. Elliott), sbaxter100@aol.com (S. Baxter).

0094-5765/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.actaastro.2011.03.007

profound effect on the human race. And, because of this, we now have initial strategies in place to cater for such a "contact" situation. We have methods for calculating the significance and impact of announcing a signal and the risk factors for replying to such a signal (Rio Scale, San Marino Scale, and First SETI Protocol [2; 3; 5; 17; 18; 19]); much has also been discussed about how we manage a post-detection announcement situation.

But what next? In the event we can prove that we have detected an extraterrestrial technology and that the signal displays intelligent-seeming structured content, we will be in a much more complex situation. No longer do issues of dissemination merely focus on announcing facts surrounding the existence of a technological "beacon"; we now find ourselves facing the complexities involved in understanding and glimpsing the intellect of the author—while the world's fears and expectations would demand immediacy of information.

The public demand to know can be summarized by a simple question: "What does the message say?" This in turn can be broken down to subquestions. "Is this a message? How large is it? How complex is it? What level of cognition produced it? Can we ever translate it—and if so, how long would it take?" To put the challenge of decipherment into context, we still have many scripts from our own antiquity that remain undeciphered, despite many serious attempts, over hundreds of years [15].

With these issues in mind, we look at the immense difficulties involved in trying to decipher the content of a signal and in communicating information on progress, while that decipherment is underway. In developing a strategy for message detection and decipherment, comparators from existing protocols for "catastrophic" and globally significant events that have high societal impact are presented as supporting rationales. Nevertheless, a post-detection scenario has very particular challenges that form its core metrics, dictating logical stages and subsequent information flow, which are significantly affected by the unknown aspects of its structure and content. To assist our capabilities in tackling such a complex task, prior research conducted on identifying structural "universals" and decipherment strategies ([9,11]), based on aspects of these computational phenomena identifiable in the constructs of language, provide essential insights into the difficulty factors each phase is likely to present.

Building on a "position" paper rationale [12], we propose an algorithmic rationale based on the previous research into signal decipherment techniques as an initial methodology for attempting to unlock the content of an extraterrestrial signal.

In terms of communication, the "DISC Quotient", in which a numerical significance factor is assigned to each of four signal analysis algorithmic steps, is a numerical method to characterize the cultural significance of the receipt of a complex and potentially decipherable signal from ETI. DISC Quotient scores would be published and updated as the analysis was underway. The purpose is to facilitate the public communication of work on any such claimed signal as such work proceeds, to provide a predictor of the work's likely outcome, and to assist in its discussion and interpretation. Analogies from human civilization are given. The most significant outcome would be a large, information-rich, and fully translatable signal from another culture: an encyclopedia, perhaps, or a Bible.

At present, the DISC Quotient is envisaged as summarized into a four-axis diagram, to be populated in the course of the signal analysis. Ultimately the four factors may be combined into a single numerical scale analogous to the Rio Scale [1], with appropriate weightings based on example cases.

This technical analysis may of course be just the first stage of the wider scholarly study of the message, as we addressed the ultimate question of the meaning and significance of a message from another culture.

3. DISC stages and numerical factors

Elliott's [12] post-detection decipherment strategy can be summarized in four stages as set out below. These four stages are steps toward the ultimate goal of translation, which is a full semantic assignment of message content, perhaps with probabilistic prioritization of alternative interpretations.

For the purposes of the DISC Quotient numerical factors δ, ω, σ, α have been assigned to each processing stage, each ranging in value from 0 to 10, 0 meaning least significance and 10 meaning maximal: DISC$=(\delta, \omega, \sigma, \alpha)$. These parameters are defined below.

1. *Characterization as a signal and first estimate of size, δ:* Given that a signal has been characterized as message-like (as opposed, for example, to image-like), a first indication of its significance is its size, perhaps in the first instance of the length of no-repeating bit stream, and later based on counts of internal components. δ is a measure of data quantity: "How big is the message?" The scores are assigned by comparison to human analogues, as defined below. Of course a comparatively short work may have a disproportionate impact; the Koran is the length of a short novel, yet has shaped human history. See examples presented in Table 1.
There is a minimum meaningful size in this context. It is estimated that a signal of complexity equivalent to human speech would require twenty thousand words ($\sim$10 exp 6 bits) to enable a full semantic analysis (of any comparable complex/information rich system).
2. *Information-theoretic analysis, ω:* ω is a measure of data quality, and a first indication of complexity and likely cognition. In early stages this can be measured using the Shannon entropy order analysis of information theory, in which any given signal can be broken down into a distribution of entropic values, a measure of internal structure and correlations, with human languages reaching a typical maximum order of 9 (see Table 2). Maximum values for different species correlate with encephalization quotients. (Note that the "chimpanzee" assessment given here is extrapolated from encephalization quotient rather than measured directly from speech analysis.)