Testing SETI Message Designs

Michael W. Busch
Division of Geological and Planetary Sciences, California Institute of Technology
Rachel M. Reddick
Department of Physics, Stanford University

Much work in SETI has focused on detecting radio broadcasts due to extraterrestrial intelligence, but there have been limited efforts to transmit messages over interstellar distances (active SETI). As a check if such messages can be interpreted once received, we conducted a blind test. One of us coded a 75-kilobit message, which the other then attempted to decipher. The decryption was accurate, supporting the message design as a general structure for communicating with aliens capable of detecting narrow-band radio transmissions.

1. Introduction

While the mechanics of transmitting radio signals across interstellar distances are well understood (e.g. NAIC 1975, Zaitsev 2006), there has been relatively little effort spent on ensuring that a transmitted message will be understandable to an alien listener.

Regardless of if active SETI is desirable, designing a message that is deliberately easy to interpret is an interesting exercise. To this end, one of us (Busch) developed a coding scheme and provided the other (Reddick) with a test message, in a blind test of decryption.

2. The Test Message

The coder based the encryption scheme on the general-purpose binary languages proposed by several authors (e.g. Freudenthal 1960, McConnell 2001), to avoid potential bias in pictorial representations (Vakoch 2000). Strings of numbers represent numbers, operators/verbs, variables/nouns, or delimiters, and are assembled into statements. The message totaled 113960 quads, but the content was on average repeated three times; the true length is ~75 kilobits.

We assumed that a watcher would be able to identify the signal as artificial due to low bandwidth, frequency modulation, and periodic Doppler shifts (e.g., SETI@home, Anderson et al. 2008). In addition, the coder assumed the watchers would be able to locate the Sun as the source of the message. The watcher was provided with a version of the message which was missing a randomly selected amount of material from the beginning (10-20% of the total) and ~2% of later quads, to represent initial detection and intermittent instrument downtime.

3. Decryption

To decrypt the message, the watcher used mostly pencil-and-paper for analysis and search-and-replace to replace deciphered blocks of the message. The watcher first noticed strings that occur in almost exactly equal numbers, serving as delimiters, then recognized the strings corresponding to = and ≠ and the notation for integers. Additional statements illustrated variables, assigning values to them, and functions for +, *, -, /, and ^. With / is the notation for floating-points, and a representation of the first nouns: e and π. As a test of arithmetic, the message includes the quadratic formula, introducing the 'or' operator.

To define a system of units, the message contains a series of formulae relating the various Planck units to each other and to dimensionless numbers (e.g. the fine structure constant). The proton, neutron, and electron are defined as nouns equal to statements providing their masses, and charges. A partial chart of the nuclides provides masses and number of protons, neutrons, and electrons, and chemical formulae follow naturally.

The next-to-last block in the message defines the Sun by mass, temperature, radius, power, and elemental composition, and refers to eight subsections of the last block, which provide masses, radii, temperatures, and a set of distances and times that follow Kepler's third law: the planets. One of the planet descriptions contains two sets of composition information: one typical of rocky terrestrial planets and one a list of gases dominated by oxygen, indicating certain chemical. Its orbit also matches the Doppler behavior of the beacon: the message defines the Earth and provides a minimalist description of terrestrial life. The watcher recognized all of this and accurately decoded it after a total of approximately twelve hours of work.

To provide additional perspectives on decryption, we provided the message independently to five undergraduates, who each spent no more than an hour attempting to decode it, again working without any pattern-recognition software. Three of the students correctly identified the delimiters, with two also identifying '=', '≠', and the notation for integers.

While we have described the decryption in a linear fashion, the message is not structured this way. Series of statements are separated from each other and repeated a varying number of times. For example, the blocks defining =, ≠, and + are given five times, while the chart of the nuclides and the chemical formulae are given only twice. Overall, the message is three times longer than the individual blocks of code. This redundancy serves the obvious purpose of allowing the message to be decrypted regardless of when the watcher started observing.

4. Comparison to Other Message Designs

As a comparison, the watcher was also provided with one of the Cosmic Call messages (Zaitsev 2006). Deciphering this message was trivial, since it is a pictorial message based on a grid. However: the information contained in our test message is roughly equivalent to that contained in the first twelve images of the 2002 Cosmic Call message. These images total 193 kilobits. In addition to avoiding possible human, constructed languages convey information more densely than purely pictorial messages (in this case, by 2.5x)

5. Conclusions

In a sharply limited number of bits, we can establish a common vocabulary and describe the solar system in considerable detail. Tentatively, those who receive it can reliably decrypt such a message. Finally, a general-purpose constructed language is more information-dense than a series of images, an advantage for beacons to be detected over large volumes.

We ask two additional questions. How much data is required for a blind decryption of more complicated ideas? And, given that there are no technical or apparent theoretical limitations, should active SETI be developed on a large scale?

Acknowledgements

We thank C.J. White and B.R. Lawrence of Caltech and three anonymous contributors for their efforts in decrypting the test message.

References:

Anderson, et al., D., 2002, **45**, 56-61.
Freudental, H., 1960.
McConnell, B.S., 2001.
NAIC, 1975, Icarus, **26**, 462-466.
Siemion, 2009, arXiv:physics/0811.3046
Vakoch, D.A., 2000, Acta Astronaut., **46**, 733-736.
Zaitsev, A., 2006, arXiv:physics/0610031v1