The Torah Code Controversy

Robert M. Haralick

Intelligent Systems Laboratory
Department of Electrical Enginnering
University of Washington
Seattle WA 98195

Abstract

In this paper we briefly explain what a Torah code is and give an example of one. We show how to calculate the probability of its occurrence in a suitably defined population of texts. We explain how it is possible to misadvertise the computed probabilities, making it seem that the probability of a Torah code is very small. We briefly discuss the controversy and suggest that new more carefully controlled experiments are needed to resolve the controversy. And we detail the protocol for a new experiment by a website reference.

1. Introduction

A Torah code is an occurrence of one or more given words spelled out by taking its successive letters at some spacing other than one in a Hebrew Torah text from which inter-word spaces and punction marks have been removed. Equal interval skips between successive letters of a word is the usual way Torah codes are found, but one can envision other skip patterns as well.

On the one hand it would seem that the formation of words formed by successive letters at equi-distant letter skip intervals (ELS) is surprising. On the other hand it would seem that since there are so many places and skip intervals to potentially form such words, that one can argue that they form just by a chance happening. And indeed one can statistically expect such formations in any kind of texts. The issue is one of knowing how to compute whether a given ELS event has high probability or low probability.

In 1994, Witztum et. al. published an article in *Statistical Science*[1] providing statistical evidence that in one suite of controlled experiments the chance probability of the Torah code patterns they found relating names of famous rabbis and dates of their births or of deaths was one in 62,500. This study used more than just the event of ELSs occurring. It attempted to measure the spatial closeness of the ELS events which did occur. Books authored by Rambsel[9] Haralick and Glazerson[3] and Novick[7] use the Torah code device to reinforce the various religious points they are trying to teach. Each tries to add weight to their religious argument by making statements about how small the probability is of the Torah code they illustrate.[8][11][10] The book by Drosnin[2] has popularized Torah codes by describing how the code for the assassination of Rabin was found a year before he was assassinated. Drosnin also shows (fallaciously) how he can make predictions using Torah codes. The book by Satinover[12] explains the ancient origin of the Torah codes and puts it in the perspective of the Jewish religious tradition. The "Discovery" seminars organized by Aish HaTorah also make use of the Torah codes to argue that their occurrece is not by chance. Therefore, they conclude that the author of the Torah was an extraordinary author.

The difficulty with the Torah code phenomena is that it is an instance of a publicly observable phenomena whose proponents argue is not a natural phenomena explainable by science. A fundamental hypothesis of science is that any publicly observable phenomena has a natural cause and effect type explanation. Therefore a considerable academic controversy[6][4][13] has arisen with the believers in science arguing that Witztum et. al.'s experiment in some way has to have been a fraud to produce such a small proability and the proponents of the Torah codes arguing for the validy of the Witztum et. al.'s experiment.[14][5]

The purpose of this article is to explain the protocols by which one can discover whether the occurrence of Torah codes is just a chance occurrence or whether it is statistically unexpected. To introduce the topic, we illustrate an example Torah code and begin the process of analyzing it to understand its statistical significance. Although our example comes from a religious context, our focus is entirely statistical and we, therefore, make no religious comments or inferences.

2. Protocols and Probabilities

In this section we analyze how to calculate the probability of a code occuring for some different protocols of searching for codes. We will first illustrate this by the Maimonides example.

Maimonides is Rabbi Moses Ben Maimon who is also known as רמבם, the Rambam, for short. He lived in Egypt in the twelfth century, 1135-1204. He was a philosopher, a physician, a halakhist, and a medical writer. He held the position of being the physician in the court of Al-Fadhil, the vizier of Egypt under Saladin. And as well, he was the head of the Jewish religous community in Cairo. Among his religious writings is the famous משנה תורה, *Mishneh Torah*, an organized compendium of the entire halakhah, the laws associated with the 613 commandments followed by observant Jews.

In the section of Exodus discussing the observance of the Passover the following Torah code for the two key words משנה תורה can be found. Each code instance has a skip interval of 50 and from the מ of משנה to the letter preceeding the ת of Torah is exactly 613 letters. This is illustrated in the code array of figure 1.

What is the probability of this occuring. To answer this question we have to put the question in the context of a protocol in which we must first define a population of texts and then describe the search experiment that is to be performed in the population of texts.

For this first code, our text population consists of texts which are all possible positional permutations of the letters in the Torah text. The Torah text has 304,805 letters and the number of possible positional permutations is 304,805!, a very large number. One such permutation is the identity permutation so the Torah text is one of the texts in the population. Our experiment is that of drawing out at random one of the texts in the population and from a pre-specified list of character positions determining whether or not the letters משנהתורה occur. Notice that the description of the experiment has two dimensions: that of selecting at random one of the texts in the population and the second the determination of whether in a pre-given list of character positions the letters משנהתורה occur.

For this text population and protocol of randomly selecting a text, we can easily compute the probability of observing the 8 letters משנהתורה in a given list of character positions. For this text population, the conditional probability of any letter in any character position given any other combination of letters in any other character position is equal to the marginal probability of the letter. So no matter what our list of letters may be and no matter what positions we specify the letters to occur in, the probability of observing the joint event is just the product of the marginal letter prob-

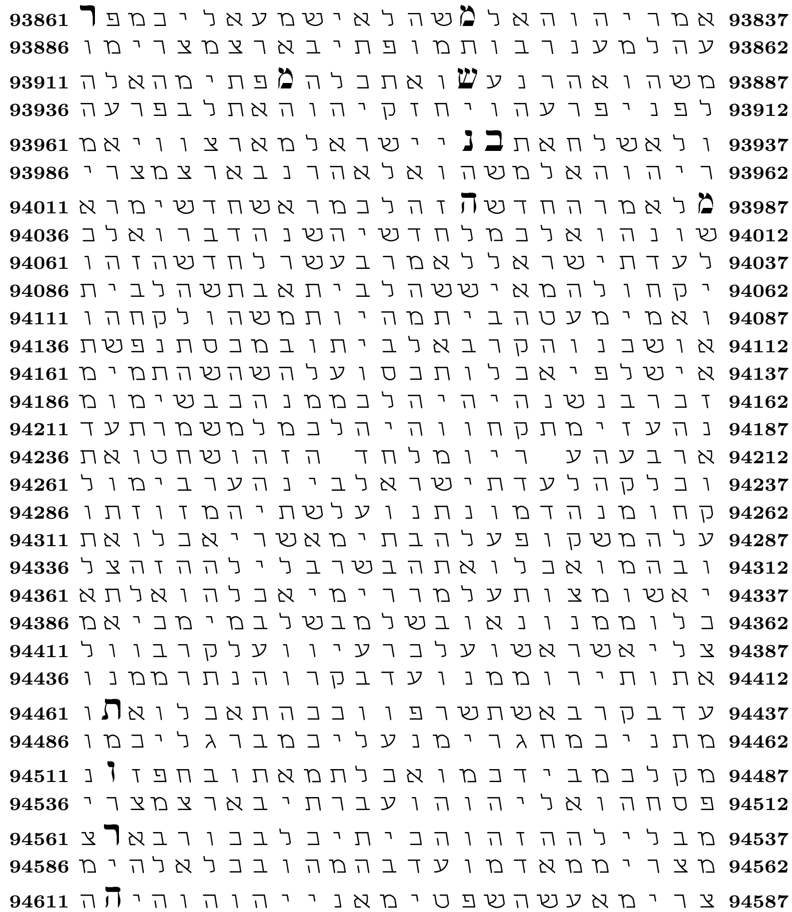

Figure 1. Code array showing the close spatial relationship between the key words רמבם, Rambam, the short nick name by which Maimonides is known, and the title of his most famous book, *Mishneh Torah* משנה תורה. The numbers on the left and the right give the text character positions for the letters in the leftmost and rightmost columns of the code array.

Letter	Probability	Letter	Probability
א	0.088774	ל	0.070766
ב	0.053624	מ	0.082314
ג	0.006919	נ	0.046351
ד	0.023070	ס	0.006014
ה	0.092045	ע	0.036909
ו	0.100106	פ	0.015764
ז	0.007211	צ	0.012998
ח	0.023585	ק	0.015403
ט	0.005918	ר	0.059464
י	0.103446	ש	0.051164
כ	0.039264	ת	0.058890

Table 1. Lists the letters and their probabilities as they occur in Torah.

abilities. Or simply stated, the event of observing any one letter in any character position is independent of observing any other letter in any other character position.

If we change the text population, we change the probability of the joint event. For example, consider the text population consisting of all permutations of the Torah text in which each permuted text has the property that each letter is part of a run of letters of its kind of length 50, with the exception of possibly each letter's last run. If we order the list of given character positions and in our ordered list each successive character position is more than 50 from the previous character position, then we have independence as before. But if the some of the successive character positions are less than 50, then independence does not hold and since the list of 8 letters משנהתורה does not have any successive repeating letters, the probability of observing them in their pre-specified positions in this new population will be zero.

Table 1 gives the letter frequency of each letter for the five books of the Torah. The probability p of observing the 8 letters משנהתורה in any given fixed set of character positions is computed as

$$
\begin{aligned}
p &= 0.082314 \times 0.051164 \times 0.046351 \times 0.092045 \times \\
&\quad 0.058890 \times 0.100106 \times 0.059464 \times 0.092045 \\
&= 5.79767 \times 10^{-10}
\end{aligned}
$$

a very small probability indeed.

Because it is so small we might think that this is an unusual event. But the way in which we might naively think that this is an unusual event could be very wrong. To understand this, we have to understand the meaning of the probability p we computed. It means this: If we were to sample one text from a population of all texts which have the same number of letters of each kind that the Torah has and if we were to designate a list of 8 particular character positions, the probability is p that we would discover the letters רה משנהתו in the designated places, precisely in this order, That is, in such an experiment, the probability is very small that we would find משנהתורה.

Suppose that we now organize a search. Suppose we select a sampling pattern of exactly a skip interval of 50 for משנה and 50 for תורה with 613 letters in between the מ of משנה and the ת of תורה. This specifies for what we are going to look: a particular pattern of 8 letters having a span of some 1+613+1+150 = 765 character positions. Now we perform the experiment. We take the Torah text having some 304,805 letters, and look in all the 304,805 - 765 +1 = 304,041 positions in which the code span of 765 character positions can be placed and look to find an occurrence of the code משנה תורה. What is the probability of observing no occurrences? What is the probability of observing one occurrence? two occurrences? three occurrence? and so on.

The probability of observing the pattern משנה תורה in any one placement is 5.79767×10^{-10}. There are 304,041 possible placements in which to observe a code that spans 765 characters. So the expected number of times or mean number of times m that we would observe the code תורה משנה in such a text is

$$m = 5.79767 \times 10^{-10} \times 304,041 = 1.76273 \times 10^{-4}$$

Assuming that the number of times that we observe the pattern is Poisson distributed, the probability q of observing the pattern k times is

$$q = \frac{\exp(-m)m^k}{k!} \tag{1}$$

To determine the probability of not observing the pattern at all we take $k = 0$ in equation 1. When $k = 0$, this probability is $\exp(-m)$, which for m near 0 is approximately $1 - m$. The probability of not observing the pattern 0 times is the probability of observing it at least once and this is $1 - (1 - m) = m$. Hence for this experiment, the probability of observing משנהתורה at least once in code is 1.76273×10^{-4}.

The probability of observing the character sequence רמבם in any given placement is

$$0.059464 \times 0.082314 \times 0.053624 \times 0.082314 = 2.16056 \times 10^{-5}$$

The number M of possible placements in a text of length N characters, searching over skip intervals from D_{min} to

D_{max} is

$$\begin{aligned} M &= \sum_{d=D_{min}}^{D_{max}} [N-(L-1)d] \qquad (2) \\ &= (D_{max}-D_{min}+1)[N-\frac{(L-1)(D_{max}+D_{min})}{2}] \end{aligned}$$

Taking $N = 765$, $D_{min} = 2$, and $D_{max} = 254$, we obtain $M = 192,786$. The expected number m of occurrences is then

$$m = 192,768 \times 2.15056 \times 10^{-5} = 4.14559$$

And the probability of observing at least one occurrence is then $1 - \exp(-m) = 0.98416$. Hence the probability of observing the letter sequence משנה at a skip interval of 50, the letter sequence of תורה at a skip interval of 50, and 613 letters between the מ of משנה and the ת of תורה, and the letter sequence רמבם, searching over skip intervals of 2 to 254, within the 765 character span the code instance of משנהתורה is $1.76273 \times 10^{-4} \times .98416 = 1.734808 \times 10^{-4}$

Now suppose that in our search for משנהתורה, we do not search just for a skip interval of 50. Suppose we search at skip intervals of say 1 to D. Then everything changes. For a skip interval of d, the span of the code instance is $615 + 3d$ character positions. The number of places such a code can be placed in a text of length 304,805 characters is $304,805-(615+3d)+1 = 304,190-3d$. Taking the sum for d being between 1 and D the total number M of possible code placements is

$$\begin{aligned} M &= \sum_{d=1}^{D} 304,190-3d \\ &= 304,190,000-3(D)(D+1)/2 \end{aligned}$$

When $D = 1000$, there results

$$\begin{aligned} M &= 304,190D-3(1000)(1001)/2 \\ &= 302,688,500 \end{aligned}$$

Hence the expected number of times m that we observe the pattern משנה תורה is

$$m = 5.79767 \times 10^{-10} \times 302,688,500 = 0.175489$$

Now the probability that we observe the pattern at least once is $1-\exp(-m) = 0.16095$, a chance of approximately one out of six times. This would certainly not be a rare event.

If the experiment is an honest experiment, meaning that the character sequence and character relationships have been specified ahead of time with complete independence from any knowledge of previous Torah code experiments, with complete independence of any previous searching in the text, then the Torah text itself can be considered as a randomly selected or arbitrarily selected text. And the probability of 0.16095 applies to it. However, it we looked first at the Torah text for the code, and we discovered it, and then we do the probability calculation, the computed probability only means: If we were to select a text at random from the population and if we were to look at this randomly selected text in all the possible placements of the code, checking each placement to see if it has in the given order the eight letters of the two key words, then the probability is 0.16095 that we will find in that text at least one Torah Code instance. But this probability does not apply to the Torah text because looking at it first disqualifies it for being a randomly selected text.

2.1. Misadvertising

The probability of observing a pattern of ELS events is relative to the protocol under which the experiment is done. Change the protocol and the probability changes. This means that it is easy to advertise an insignificant result as a signiicant result by finding a Torah code in an experiment with one kind of protocol but advertising that it was found with a different protocol.[13]

This kind of misadvertising can take three forms. The first form is to do experiments with many key word sets and select out those key word sets having smaller associated probabilities. Then devise a new experiment using a combined key word set consisting of those key word sets having the smaller associated probabilities. This new experiment is publicly reported. Just to show such an effect. Bar-Natan and McKay[6] did a parlor room Witztum like Torah code experiment with a modern Hebrew translation of *War and Peace*. The probability of the publicly reported expriment was very small, almost as small as in the Witztum et. al. experiment. They concluded that whatever the phenomena is that occurred in the Torah text, it was also present in the *War and Peace* text. Thus Bar-Natan and McKay attempt to make a mockery of the Witztum et. al. work.

3. Settling the Controversy

The current situation is one in which the skeptics can hypothesize a protocol involving one or more prior experiments done on the sly, by which the small probability results observed in the publicly declared experiment are expected, explainable without miracle, and therefore not statistically signficant. But the situation the skeptics describe is not consistent with the exprimenters repeatedly made assertions that there were no experiments on the sly and there was no peeking ahead.

One argument counters the other. In this case, what is the appropriate action for a scientist? What does a scientist do when there occur inconsistencies between observations, or between observations and theory? To discover the truth, the scientist designs a new experiment, an experiment which is more carefully controlled for extraneous effects, one which is publicly open, an experiment in which the instrumentation is more accurate, a data analysis protocol in which the statistics computed are more robust and have smaller expected variances etc. This is what is required here.

To help settle the controversy new experiments are needed. A full statement of the protocols for a new experiment can be found in the web site http://www.george.ee.washington.edu. The experiment in certain respects is patterned after that of Witztum et. al.[1] to test the hypothesis that ELSs of appellations by which famous rabbis are known and ELSs of their birth or death dates are spatially closer together in the Torah text for more rabbis than would be expected by chance. The statistical analysis methodology is called the best star team methodology and can also be found in the above mentioned website.

References

[1] E. R. Doron Witztum and Y. Rosenberg. Equidistant letter sequences in the book of genesis. *Statistical Science*, 9:429–438, 1994.

[2] M. Drosnin. *The Bible Code*. Simon and Schuster New York, 1997.

[3] R. Haralick and M. Glazerson. *Torah Codes and Israel Today*. Lev Eliyahu Jerusalem, 1996.

[4] G. Kalai. On the paper of witztum, rips, and rosenberg on equidistant letter sequences in the book of genesis. draft, August 21, 1997.

[5] D. Mechanic. Torah codes and kiruv rechokim. *Jewish Action*, Spring:33–37, 1998.

[6] D. B. Natan and B. McKay. Draft paper. *internet website*.

[7] N. A. Novick. *Fascinating Torah Prophecies Currently Unfolding*. Netzach Yisrael Publications new York, 1997.

[8] N. A. Novick. *Fascinating Torah Prophecies Currently Unfolding*. Netzach Yisrael Publications, New York, 1997, p 148.

[9] Y. Rambsel. *Yeshua*. Frontier Research Publications, Ontario, 1997.

[10] Y. Rambsel. *Yeshua*. Frontier Research Publications, Ontario, 1997 p1.

[11] Y. Rambsel. *Yeshua*. Frontier Research Publications, Ontario, 1997, p95.

[12] J. Satinover. *Cracking the Bible Code*. William Morrow and Company New York, 1997.

[13] B. Simon. A skeptical look at the torah codes. *Jewish Action*, Spring:17–24, 1998.

[14] D. Witztum. The seal of god is truth. *Jewish Action*, Spring:25–32, 1998.