Torah Codes: New Experimental Protocols

Robert M. Haralick

Intelligent Systems Laboratory
Department of Electrical Engineering
University of Washington
Seattle, WA 98195

Abstract. Because of its unexpected nature, finding words in as equidistant letter sequences (Torah codes) in a text may appear to be interesting. However, there is a significant probability that they occur by chance. In this paper we discuss a repeatable and objective methodology for defining a priori related pairs of key words for use in an experiment and we define a testing methodology for testing the hypothesis of whether such related equidistant letter sequences found in the Torah text are more spatially close in the text than expected by chance.

1 Introduction

A Torah code is an occurrence of one or more given words spelled out by taking its successive letters, at some distance other than one from each other, 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 equal letter skip intervals is surprising. On the other hand it would seem that since there are so many ways to potentially form such words, that one can argue that they form just by a chance happening. The question is whether the observed phenomena is just a chance phenomena.

Books authored by Rambsel[9] and Novick[8] use the Torah code device to reinforce the religious point they are trying to teach. The "Discovery" seminars organized by Aish HaTorah also make use of the Torah codes to argue that they happen not by chance. Therefore, they conclude that the author of the Torah was an extraordinary author. Witztum et. al. published an article in *Statistical Science* 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 deaths was one in 62,500. In the past year there have been two books about Torah codes[10, 7] and many articles debating the findings[11]. In this paper, we briefly describe the Witztum et. al. experiment, and the new series of experiments we propose to carry out to test some hypotheses about whether the Torah codes are real or not.

2 Protocols and Probabilities

We illustrate the issue of protocols and probabilities by a famous Torah code example involving Rabbi Moses Ben Maimon who is also known as Maimonides. 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 ELS has a skip interval of 50 and from the מ of משנה to the letter preceeding the ת of Torah is exactly 613 letters, the 613 corresponding to the 613 commandments. This is illustrated in the code array of figure 1.

Having observed this phenomena in a given text, we could ask what is the probability of this occuring. This question is actually more complicated than it seems on the surface. First we must understand that probabilities in this context are probabilties relative to an experiment. And we must be very careful in specifying the experiment.

One possible experiment is that we were exploring the given text in a non-structured way and notice an equidistant letter sequence (ELS) pattern. Then we ask the question of what would the probability be of observing this ELS pattern in a text from a suitably defined text population. And the probability answer, p, we obtain is relative to an experiment that we would do on the text population. The meaning of the probability p is that if we were to randomly sample a text from the population and see if the text contained the ELS pattern, the probability of the randomly sampled text containing the ELS pattern would be p. However, this probability p has nothing to do with our original text, even if the original text is a member of the text population. The reason it has nothing to do with the original text is that the original text was examined first and then the probability question was asked. So it is a probability after the fact. Probabilities obtained after the fact are meaningful relative to a future experiment that might be done on a randomly sampled text, but not relevant to a past experiment that has been done on a given text already sampled from the population and examined.

On the other hand, if we first define the ELS pattern and then ask the question what is the probability of finding the ELS pattern in some text of a suitably defined text population of which our given text is a member, then the probability answer is applicable to any text in the population including our given text, because we have not yet explored our given text.

From this analysis, we understand that since we have not been told the experimental protocol of the Maimonides example, we cannot say that any probability we compute pertains to the Torah text in which the Torah code ELS pattern was found.

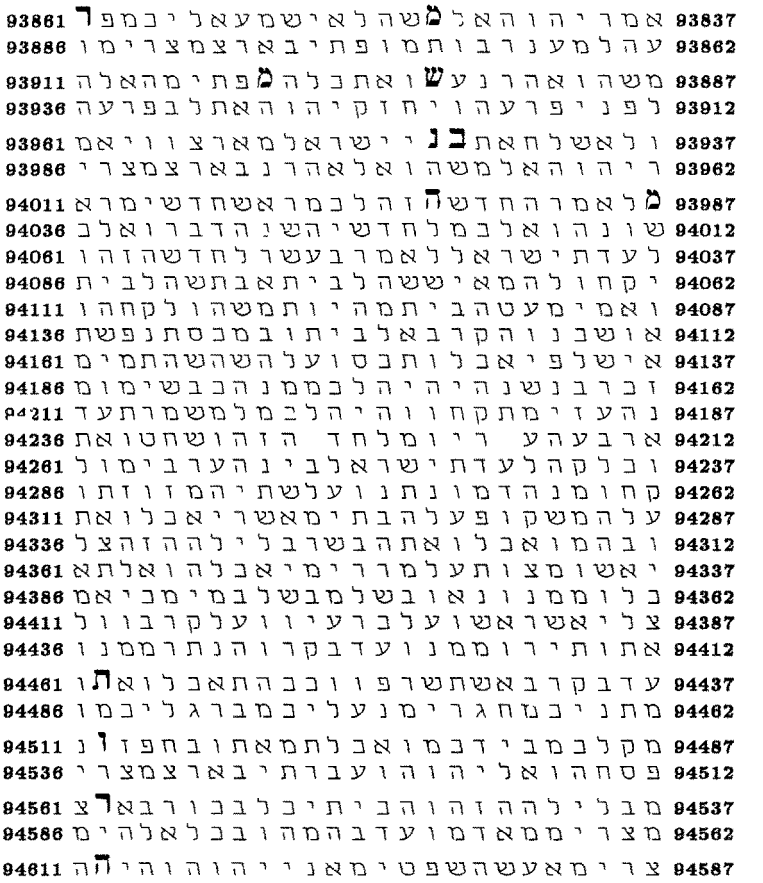

Fig. 1. Code array showing the close spatial relationship between the key words רמבם, Rambam, the short nick name by which Maimonides is known, and the title of his most famous book, *Mishneh Torah* משנה תורה. The numbers on the left and the right give the text character positions for the letters in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

3 The Witztum et. al. Experiment

Witztum .et. al.[6] did the following experiment. They looked in the *Encyclopedia of Great Men In Israel*, selecting all men who had between one and a half and three columns of text written about them. Each rabbi had one name and possibly some other appellations by which the rabbi was known. Each rabbi had a date consisting of day and month associated with his birth date or death date. Each

date had two or three ways of writing it.[1] A query then consists of a paired name set and date set, the name set consisting of all the appellations of the rabbi and the date set consisting of all ways of writing the dates that have more than four characters. They devised four different statistical measures related to the compactness, the spatial closeness, with which a name and date ELS pair are associated in the text. For details about the compactness measure, see [6]. For our purposes of description we need only understand that compactness is what they considered to be a suitably normalized measure of how close the approximately ten smallest skip interval ELSs of each appellation and date pair are in the Genesis text. A multiple query run then produced four measures of compactness for each appellation date pair for each rabbi. Then for each compactness type, they combined all the resulting compactness numbers to obtain one number for the experiment.

They then repeated the experiment 999,999 more times, each time randomly permuting the name set date set associations. For each of these 999,999 times, four measures of compactness were produced. Then the value of compactness produced by the correct name date association is compared against the value of compactness produced by the 999,999 random name-date associations. And the number of times that a random association produced a more compact value was counted. The results of this counting for their four compactness measures was the four values: 453, 5, 570, and 4. Using the Bonferoni inequality, the p-value (significance level) of this experiment is no more than

$$\frac{4\ min\{453, 5, 570, 4\}}{1,000,000} = \frac{16}{1,000,000} = \frac{1}{62,500}.$$

They, therefore, concluded that in Genesis, the proximity of equidistant letter sequences for the names and dates selected was not due to chance.

The essential elements of the protcol they used is that by specifying the list of rabbis to come from a published encyclopedia using a simple selection method, the selection of the rabbis must be considered as an aprior selection. This list of rabbis was given to Prof Havlin, of Bar Ilan University, to provide the associated list of appellations and dates for each rabbi. Prof. Havlin is an historical scholar and has expertise in this area. Since Prof. Havlin did not do any Torah code experiments, his preparation of the lists counts as apriori to the experiment.

Unfortunately, the experiment itself has a difficulty relative to the population. It uses a population of monkey queries – mixed up appellation and date sets. So the result of the experiment may say something directly about the unusualness of the correct pairing of appellation and date sets in a population of random pairing of appellation and date sets, but it does not say something directly about the Torah text which is how the naive person would understand

[1] There are four ways of writing a day and a month in Hebrew, three of which were used in the Witztum experiment. For technical reasons of how Witztum et. al. chose to normalize the observed compactness between appellationa and date, the normalization could not be done with dates having less than five characters. Hence some ways of writing dates had to be excluded.

the experiment. But perhaps even more importantly, for the ranking method the experiment used to estimate the chance probability, all queries in the monkey query population have to have the same probability structure, a symmetry condition. But because the appellation and date sets for each rabbi have different sizes, each mixing up of date sets with the appellation sets produces a different number of appellation date pairs over the set of rabbis. This makes the total compactness value associated with each query come from a different distribution. There are other criticisms that can be made of the Witztum experiment, but there is not space here to detail them.

The real problem with the experiment is the questions rasied by the skeptics. They ask: is the appellation selection objective? Would another person produce the same list of appellations. For the degree to which subjective elements enter the appellation gathering, is the degree to which the subjective elements can tilt the experiment to a small probability. Thus it is important that the appellation gathering be repeatable by any another person. The repeatability guarantees that there are no subjective elements that can tilt the experimental result. Furthermore, the skeptics ask: is everything a priori? What guarantees are there in the protocol that the list of appellations is truly apriori. How do we know that there was not a larger list of appellations and the experiment was actually run twice. And after examining the results of the first experiment, an experiment done on the sly, appellations that would contribute to a large chance probability were removed. Then when the second experiment was run, the resulting probability produced by the experiment would be expected to be small. And it is this small probability that was reported in the Statistical Science paper of Witztum et. al. Under the implied protocol of the skeptics, this small probability had in fact a large probability of occuring in the second experiment and is therefore not statistically significant.

Of course, the original experimenters maintain that they are honest and everything was done in accordance with the specified protocol and the selection of appellations is objective.

Rather than arguing for or against, we will follow the scientific experimental method. To bring out the truth, science proceeds by repeating experiments refining controls and improving the data analysis protocols. It is in this spirit that the rest of this paper is written.

4 The New Experiments

The new experiments we propose to do and described here will be done in three phases. The first phase or first experiment, involving the rabbis of Table 1 of Witztum, Rips, and Rosenberg[6], is a pilot phase to iron out any kinks in the methodology. The second phase, involving the rabbis of Table 1 and 2 of Witztum, Rips and Rosenberg[6] to is provide assurance that the kinks have been ironed out. The final phase will involve all rabbis of Margalioth[4] having 50 lines or more text.

Each experiment will involve a population of texts. We take this population to be random within chapter word permuted Genesis texts. Such a text population has the advantage that the letter distribution and word distribution and word and letter clumping, chapter by chapter, for each text in the population will be similar to that in the Genesis text. Of course most of the texts in the population will not be linguistically meaningful.

For each appellation or date key word, the largest skip interval for searching can be unlimited or be set so that the expected number of ELSs found will be 5 or 10. So each experiment will have a parameter for the expected number of ELSs to be found. This parameter will in turn set largest skip interval limits, word by word, for each key word.

There are two easily expressed and understood compactness measures associated with the ELSs of an (appellation,date) pair. The 1D measure is the length of the shortest text segment that contains the span of at least one ELS for the appellation key word and at least one ELS for the date key word. The 2D measure is the number of characters in the smallest sized window in a code array of the shortest length text segment that contains the span of at least one ELS for the appellation key word and at least one ELS for the date key word. In addition to these compactness measures, the un-normalized measure Ω used in the original experiment will also be used. So each experiment will have a parameter indicating which combinations of compactness measures will be used.

There are a number of hypothesis that could be tested. The null hypothesis is that the ELS compactness values we observe are just due to chance. This null hypothesis can be tested against a variety of alternative hypotheses.

1. for all rabbis all appellation date pairs are encoded
2. for all rabbis some appellation date pairs are encoded
3. for some rabbis, more than would be expected, some, more than expected, appellation date pairs are encoded
4. some, more than expected, appellation date pairs are encoded among all the ELS pairs from all the rabbis

If some appellation date pairs are encoded, it implies that there are some rabbis who have some of their appellation date pairs encoded. But the alternative hypothesis (3) says more. There are two levels of chance. One level of chance gives some appellation date pair ELSs high compactness and there is another level of chance that gives some number of rabbis some appellation date pair ELSs that have high compactness. So under the null hypothesis of no Torah code effect, there is a distribution for the number of high compactness ELS pairs and there is a distribution for the number of rabbis having some high compactness ELS pairs. This difference can be concretely understood by assuming for the moment that we have in hand some N ELS pairs that have been determined to have statistically significant high compactness. One possibility is that they could all be ELS pairs associated with one rabbi or they could be distributed some to each rabbi. So it is possible for the null hypothesis to be rejected when tested against alternative hypothesis (3) but not to be rejected when testing against alternative hypothesis (4), and of course vica versa.

In our experiment, the null hypothesis to be tested is that the observed compactness of the appellation date ELS pairs are as expected just due to chance formations. For the alternative hypothesis, we choose either alternative hypothesis (3) or alternative hypothesis (4). Alternative (3) is that the number of highly compact ELS appellation date pairs is significantly higher than expected by chance. Alternative (4) is that the number of rabbis having highly compact ELS appellation date pairs is significantly higher than expected by chance.

5 Appellation Selection

This section first discusses an objective appellation and date gathering methodology. The raw lists gathered with this methodology can be found in the Torah code web page http://george.ee.washington.edu. Then we discuss the preparation methodology for using these raw lists to generate the appellation and date lists required for the experiment. Finally for the rabbis of Table 1 of [6] we give the appellation and date lists to be used in the first phase of the new experiment.

To eliminate the possible criticism that any possible subjective element entered the preparation of the appellations to tilt the experimental results, the methodology must be as mechanical, consistent, reasonable, and replicable as possible. For this purpose we use four Hebrew biographical collections[2, 3, 4, 5]. We also use an English significant date collection[1] as an additional source for death dates. The appellations were taken from the headings of each entry, which are set centered usually in a bigger or bolder font, and if they contain names, the words immediately after the heading.

To make for consistency through the different stylistic variations even in the same collection, if certain key words occurred in the four lines after the heading, the appellations after these key words were taken as well. The key words used to indicate that an appellation follows are מכונה, מפורסם בכינויו, הנקרא, מכונה קיצור, מפורסם על שם ספרו, and מחבר.

The dates listed in our raw table are exactly as given in the text. Sometimes instead of specifying the first of the month, the specification is given as ר"ח or as spelled out as ראש חודש. In addition, some dates were given relative to a holiday, such as בערב שבת חנוכה, ג דחול המועד סוכות, and שבועות ב. These too are listed as they appear in the biographical collections. In some years, the month of אדר occurs twice. The collections either show this by writing אדר ב or אדר שני. These too are listed exactly the way they occur in the collections.

The raw table of appellation and dates gathered from the collections has to be organized and redundancies removed to prepare the lists that will be used in the actual experiment. For the purposes of the experiment we have divided appellations into different possibilities: the given name, the family name, the common name, a title name, a name associated with a city, a name associated with a book, a name associated with the father. Each listing for a category has one name. Since there may be multiple given names or family names, these categories may appear listed multiple times. Also the spelling of family names

can vary among the sources we consulted. Every spelling variation in the sources is listed. So for this reason a category may appear multiple times. In the case of a 2 character last name, if this last name is preceeded by the titular designation רבנו, then we group it together with the last name. Otherwise we group it with the preceeding or following appellation.

A common name, often made up of the first letters of the rabbi's name, like rabbi Moses Isserles, רבי משה איסרלש, whose common name is Rama, רמא, sometimes also is referenced in our sources with the prefix ה, meaning *the*, הרמא. When this happens, we let the common name category be listed twice, once without the ה prefix and once with the ה prefix. There are also some common names which have only two letters. In this case we uniformly add the prefix ה to make it a three letter name.

A book title may have multiple words. We have kept all the words of a book name in the same category. Sometimes a book title has associated with it the word בעל. In this case we list the book title appellation with and without the word בעל. Likewise in case of a name that has the word בעל, we list the name with and without the word בעל.

The city name category often has the prefix מ, meaning *from*. The city name is listed as it appears in the entry of the sources. If it does not have a מ prefix, then we list it without the prefix. If it does have the מ prefix, then we list it with the prefix. If in one source it does and in another source it does not, then we list the category twice, once with the מ and once without the מ prefix. Also there can be some variation in the spelling of the city name. Again, we provide multiple listings of the city category to include all spelling variations that occur in our sources.

The last category is the father's name. The father's name could be a single appellation or a multiple appellation. In case of a multiple appellation, we list all the appellations of the father's name in this category. Also we prefix the father's name once by the prefix בר and once by the prefix בן, both meaning *son of*. Sometimes it was not clear from a source whether an appellation was an appellation for the rabbi or for the rabbi's father. In such cases we were guided by Margalioth[4] who lists the father's name in parentheses. Sometimes the father's name is given in our sources as a first and last name. Although we were tempted to take the last name of the father and give it as an appellation to the rabbi, we grouped the father's last name with the first name in the category father. Sometimes the name of a father or grandfather becomes part of a name, such as אבן עזרא. In this case we group אבן together with עזרא.

The death dates were not always consistent among the sources. Sometimes we saw what appeared to be type setting errors. And in once case we are sure we found a rabbi mixup in our sources. Not all the sources listed a month and day death date, although almost all listed the year of death. Therefore, our rule was to use any death date which occurred a majority of times a death date of month and day was given, ignoring any listing that gave a date relative to a holiday or as Rosh Chodesh of a month. This leaves the possibility that there might only be four sources that give a month and day death date, two of which

give one date and two of which give another date. And indeed that did happen in the case of the rabbi mixup. There are actually two different rabbis whose name is Gershon Asheknazi. And both rabbis wrote a book with the same title: עבודת הגרשוני, *Avodat Gershoni.* One died in the Jewish year 5453 and one died in the Jewish year 5466. Two sources reference one of these rabbis and two sources reference the other. To be consistent with the list of rabbis in the Statistical Science article[6], we selected in the edited list the rabbi who died in 5453.

Our rule also leaves open the possibility that none of the sources which give a month and day death date agree. In this case our rule was to delete the rabbi from the list. This however never happened. Finally, there is the issue of the spelling of the month of Cheshvon, חשון. Often this spelling is not used because it has a meaning of an accounting, a judgement. So the alternate spelling is מרחשון. And indeed this is the spelling that most of our sources preferred and so we use it.

5.1 Declaration

It is appropriate to say that with only one exception, none of the subjective choices for the rules used to gather the list of appellations and dates and none of the subjective choices used to create the edited lists were based on or influenced by any knowledge of any computer Torah code experiments done by others or myself. The exception to this is Gerson Ashkenazi who appears to be the name of two different rabbis. As we previously mentioned, for this rabbi we did consult with the Statistical Science article to select the rabbi and date that was consistent with the Statistical Science experiment. Throughout this period of gathering the data and forming the final lists, I have only done computer Torah code runs in the process of checking out programming changes that I have been making to the computer Torah code programs that I will be using in the experiment.

6 Best Star Team Analysis Methodology

The "best star team methodology" is a statistical testing methodology for testing whether the compactness of equidistant letter sequences (ELS) pairs in a text associated with the kind of experiment done by Witztum et.al. is due to chance or whether they in fact have statistically significant smaller (more compact) values. The statistics question is how to define the meaning of a larger than expected left tail for a multivariate distribution of the compactness values. In this section we first give a mathematical description of the test and then describe how the test is carried out in a Monte Carlo mode.

Appellation and Death Date List

Type	Appellation	Death Date
1	Abraham Ben David of Posquieres	
Given name	אברהם	כו כסלו
Title	הראבד	בכו כסלו
Title	הראבד השלישי	כו בכסלו
Book	בעל ההשגות	בכו בכסלו
Book	ההשגות	
Book	בעל השגות	
Book	השגות	
City	מפושקירא	
City	מפושקיירא	
City	מפאסקיירי	
Father	בן דוד	
Father	בר דוד	
2	Abraham Maimon HaNagid	
Given name	אברהם	יח כסלו
Common Name	החסיד	ביח כסלו
Common Name	מימוני	יח בכסלו
Father	בר משה	ביח בכסלו
Father	בן משה	
Father	בן הרמבם	
Father	בר הרמבם	
Father	בן משה בן מימון	
Father	בר משה בן מימון	
3	Abraham Ibn Ezra	
Given Name	אברהם	א אדר
Given Name	אבן עזרא	בא אדר
Title	ראבע	א באדר
Father	בר מאיר	בא באדר
Father	בן מאיר	

Type	Appellation	Death Date
4	Elijah Ben Asher (Bahur) Levita	
Given Name	אליהו	ו שבט
Given Name	בחור	בו שבט
Family Name	הלוי	ו בשבט
Family Name	אשכנזי	בו בשבט
Father	בר אשר	
Father	בן אשר	
Book	באל התשבי	
Book	התשבי	
Book	הבחור	
5	Elijah Ben Solomon Zalman	
	Gaon of Vilna	
Given Name	אליהו	יט תשרי
Family Name	זלמן	ביט תשרי
Title	הגרא	יט בתשרי
Common Name	הגאון	ביט בתשרי
City	מווילנה	
City	מווילנא	
City	מוילנא	
Father	בן שלמה	
Father	בר שלמה	
6	Gershon Ashkenazi of Metz	
Given Name	גרשון	י אדר
Family Name	אשכנזי	בי אדר
Family Name	אוליף	י באדר
Book Name	בעל עבודת הגרשוני	בי באדר
Book Name	עבודת הגרשוני	
Father	בר יצחק	
Father	בן יצחק	
7	David Gans	
Given Name	דוד	ח אלול
Family Name	גאנז	בח אלול
Family Name	אוזא	ח באלול
Book	בעל צמח דוד	בח באלול
Book	צמח דוד	
City	מפראג	
Father	בר שלמה	
Father	בן שלמה	

Type	Appellation	Death Date
8	David Ben Samuel Halevi Taz	
Given Name	דוד	כו שבט
Family Name	הלוי	בכו שבט
Common Name	הטז	כו בשבט
Book	טורי זהב	בכו בשבט
Book	בעל טורי זהב	
Father	בן שמואל	
Father	בר שמואל	
9	Hayyim Ben Moses Attar Or HaHayyim	
Given Name	חיים	טו תמוז
Family Name	בן עטר	בטו תמוז
Book	בעל אור חיים	טו בתמוז
Book	אור חיים	בטו בתמוז
Father	בר משה	
Father	בן משה	
10	Judah Ben Asher Son of Rosh	
Given Name	יהודה	יז תמוז
City	מטוליטולא	ביז תמוז
Father	בר אשר	יז בתמוז
Father	בן אשר	ביז בתמוז
Father	בן הראש	
Father	בר הראש	
11	Judah Ben Samuel HeHasid	
Given Name	יהודה	יג אדר
Common Name	החסיד	ביג אדר
Father	בר שמואל	יג באדר
Father	בן שמואל	ביג באדר
12	Judah Ben Bezalel Maharal of Prague	
Given Name	יהודה	יח אלול
Family Name	ליווא	ביח אלול
Family Name	לוואי	יח באלול
Common Name	מהרל	ביח באלול
City	מפראג	
Father	בר בצלאל	
Father	בן בצלאל	

Type	Appellation	Death Date
13	Jonathan Eybeschuetz of Prague	
Given Name	יהונתן	כא אלול
Family Name	אייבשיץ	בכא אלול
City Name	מפראג	כא באלול
Book	כרתי ופלתי	
Father	בר נתן נטע	בכא באלול
Father	בן נתן נטע	
14	Joshua Hoeschel	
Given Name	יהושע	כז אב
Family Name	חריף	בכז אב
City	מקראקא	כז באב
City	מקראקוב	בכז באב
City	מקרקא	
Book	מגיני שלמה	
Father	בר יוסף	
Father	בן יוסף	
15	Joshua Falk Ben Alexander Katz	
Given Name	יהושע	יט ניסן
Given Name	יושע	ביט ניסן
Family Name	ולק	יט בניסן
Family Name	פלק	ביט בניסן
Family Name	כץ	
Family Name	הכהן	
Common Name	סמע	
Book	ספר מאירת עינים	
Father	בר אלכסנדר	
Father	בן אלכסנדר	
16	Joel Sirkes (Bach)	
Given Name	יואל	כ אדר
Family Name	סירקיש	בכ אדר
Family Name	סירקים	כ באדר
Title	אבד	בכ באדר
Common Name	בח	
Common Name	הבח	
Common Name	בית חדש	
City	קראקא	
Father	בר שמואל יפה	
Father	בן שמואל יפה	

Type	Appellation	Death Date
17	Yom Tov Lipmann Heller	
Given Name	יום טוב	ו אלול
Family Name	ליפמן	בו אלול
Family Name	ליפאמן	ו באלול
Family Name	סגל	בו באלול
Family Name	הלר	
Family Name	העללער	
Book	תוספות יום טוב	
Father	בר נתן הלוי	
Father	בן נתן הלוי	
18	Jonah Ben Abraham Gerondi	
Given Name	יונה	ח מרחשון
Common Name	החסיד	בח מרחשון
City	מגירונדי	ח במרחשון
City	גירונדי	בח במרחשון
City	מגירונדא	
Father	בר אברהם	
Father	בן אברהם	
19	Joseph Caro	
Given Name	יוסף	יג ניסן
Family Name	קארו	ביג ניסן
Book	בית יוסף	יג בניסן
Book	הבית יוסף	ביג בניסן
Book	בעל השלחן ערוך	
Book	השלחן ערוך	
Father	אפרים	
20	Ezekiel Landau of Prague	
Given Name	יחזקאל	יז איר
Family Name	לנדא	ביז איר
Family Name	סגל	יז באיר
Book	נודע ביהודה	ביז באיר
City	מפראג	
Father	בר יהודה הלוי	
Father	בן יהודה הלוי	

Type	Appellation	Death Date
21	Jacob Joshua Falk	
Given Name	יעקב	יד שבט
Given Name	יהושע	ביד שבט
Family Name	פלק	יד בשבט
Book	פני יהושע	ביד בשבט
Father	צבי הירש	
22	Jacob Ben Meir (Tam)	
Given Name	יעקב	ד תמוז
Common Name	רבנו תם	בד תמוז
Common Name	רת	ד בתמוז
Book ?	בעל התוס	בד בתמוז
Father	בר מאיר	
Father	בן מאיר	
23	Isaac Alfasi (Rif)	
Given Name	יצחק	י איר
Family Name	אלפסי	בי איר
Common Name	ריף	י באיר
Common Name	הריף	בי באיר
Father	בר יעקב הכהן	
Father	בן יעקב הכהן	
24	Israel Ben Eliezer Ba'al Shem Tov (Besht)	
Given Name	ישראל	ז ניסן
Common Name	בעל שם טוב	בז ניסן
Common Name	בעשט	ז בניסן
Father	בר אליעזר	בז בניסן
Father	בן אליעזר	
25	Meir Ben Baruch of Rothenburg Maharam	
Given Name	מאיר	יט איר
Common Name	מהרם	ביט איר
City	מרוטנברג	יט באיר
Father	בר ברוך	ביט באיר
Father	בן ברוך	

Type	Appellation	Death Date
26	Mordecai Ben Abraham Jaffe	
Given Name	מרדכי	ג אדר ב
Family Name	יפה	בג אדר ב
Title	אבד	ג באדר ב
Book	הלבושים	בג באדר ב
Book	לבושים	ג אדר שני
City	פוזן	בג אדר שני
Father	בר אברהם	ג באדר שני
Father	בן אברהם	בג באדר שני
27	Moses Isserles (Rama)	
Given Name	משה	יח איר
Family Name	איסרלש	ביח איר
Family Name	איסרליש	יח באיר
Common Name	הרמא	ביח באיר
Common Name	רמא	
Father	בר ישראל	
Father	בן ישראל	
28	Moses Hayyim Luzatto	
Given Name	משה	כו איר
Given Name	חיים	בכו איר
Family Name	לוצאטו	כו באיר
Common Name	רמחל	בכו באיר
Book	מסלת ישרים	
Father	בר יעקב חי	
Father	בן יעקב חי	
29	Mose Ben Maimon Rambam, Maimonides	
Given Name	משה	כ טבת
Common Name	הרמבם	בכ טבת
Title	הגדול	כ בטבת
Father	מימון	בכ בטבת
30	Zevi Hirsch Ben Ashkenazi	
Given Name	צבי	א איר
Family Name	הירש	בא איר
Family Name	אשכנזי	א באיר
Family Name	מלבוב	בא באיר
Book	החכם צבי	
Book	חכם צבי	
Father	בר יעקב	
Father	בן יעקב	

Type	Appellation	Death Date
31	Shabbetai Ben Meir HaKohen (Shakh)	
Given Name	שבתי	א אדר
Family Name	הכהן	בא אדר
Common Name	השך	א באדר
Father	מאיר	בא באדר
32	Shelomoh Yizhaki (Rashi)	
Given Name	שלמה	כט תמוז
Family Name	יצחקי	בכט תמוז
Common Name	רשי	כט בתמוז
Title	הגדול	בכט בתמוז
Father	יצחק	
33	Solomon Luria (Maharshal)	
Given Name	שלמה	יב כסלו
Family Name	לוריא	ביב כסלו
Family Name	אשכנזי	יב בכסלו
Common Name	מהרשל	ביב בכסלו
Common Name	המהרשל	
Father	בר יחיאל	
Father	בן יחיאל	
34	Samuel Eliezer Ben Judah Halevi Edels Maharsha	
Given Name	שמואל	ה כסלו
Given Name	אליעזר	בה כסלו
Family Name	אידליש	ה בכסלו
Family Name	איידלס	בה בכסלו
Family Name	אידלס	
Common Name	מהרשא	
Common Name	המהרשא	
Father	בר יהודה הלוי	
Father	בן יהודה הלוי	

6.1 The Experiment Overview

The best star team methodology is appropriate to anaylze experiments which have the following essential characteristics:

1. A set of personalities.
2. A set of (appellation,date) pair key words associated with each personality.

3. A Torah text and an associated population of monkey texts. The population monkey texts can be, for example, permuted word within chapter Torah texts.
4. A random sampling of texts from the population, one of the sampled texts being the Torah text.
5. At least one kind of a compactness measure which when given a pair of ELSs corresponding to a pair of appellation and date key words produces a value of the compactness of the pair of ELSs.
6. A decoder which when given an (appellation,date,compactness type) triple first computes for each text sampled from the population the ELSs of the given appellation and date. If one or the other or both have no ELSs, then it produces a special value *. If both have ELSs then it produces a number associated with the smallest compactness of the given compactness type among all (appellation,date) ELS pairs found for the given (appellation,date) pair.

6.2 Abstract Description

Let $X_1, ..., X_N$ be N random variables whose cumulative distribution function is known $F(X_1, ..., X_N)$. Independence among the $X_1, \ldots, X_N$ is not assumed and the X_n are not assumed to be identically distributed.

Define $G_n(X_n)$ to be the marginal cumulative distribution function for X_n. Let $Y_n = G_n(X_n)$. This normalizes the random variables to uniforms. This makes them probability (scale) comparable.

Let $Y_{(1)}, \ldots, Y_{(N)}$ be the order statistics of $Y_1, \ldots, Y_N$. The smallest value among them is the raw value of a star team of one member. The average of the two smallest values among them is the raw value of a star team of two members and so on. Define

$$Z_n = \frac{1}{n} \sum_{i=1}^{n} Y_(i)$$

Each Z_n is then the raw value of the star team of n members. It must be that $Z_1 <= Z_2 <= ... <= Z_N$. Also notice that since $Y_(i)$ are scale comparable, the different random variables that might be in the sum from experiment to experiment nevertheless produce comparable averages.

Let H_n be the marginal distribution function for Z_n. Let $Q_n = H_n(Z_n)$. This normalizes the random variables to uniforms and gives a probability interpretation to Q_n. Q_n is the normalized score for the star team of n members.

Let $S = \min\{Q_1, ...Q_N\}$. S represents the smallest normalized value of the star teams. Hence it can be thought of as the normalized value for the best star team. Let R be the marginal cumulative distribution function for S. Then $R(S)$ is the probability of there arising a multivariate left tail as small or smaller than observed.

In the Torah code application the function F associated with the null hypothesis that the left tail is small is not known. However, it is possible to define a suitable population consistent with the null hypothesis from which values of

$X_1, ..., X_N$ can be sampled. So by Monte Carlo simulation we may obtain thousands of independent trials, each of $X_1, ..., X_N$. To get the Y_n, we can use the empirically determined G_n. Having it, we can determine for each trial each Z_n. Having it, we can use the empirically determined H_n to define the Q_n. Having the Q_n defined, we can for each trial determine the smallest among them. This defines the S for each trial. And having the S for each trial, including the first trial, which used the Torah text and whose value is S_0, we can use the empirically determined distribution R of S to determine (estimate) the $R(S_0)$ for the initial $X_1, ...X_N$ we observed. We reject the null hypothesis in favor of the hypothesis that the multivariate left tail is fatter than expected at the p significance level if $R(S_0) < p$.

6.3 The Monte Carlo Mode

Our analysis of whether the results are significantly different from chance will treat the Torah text and each monkey text exactly the same: the star team consisting of the champion (appellation,date,compactness type) triples for each will be used. Our analysis will compare the ratings of these star teams to test the null hypothesis against the alternative that there are more rabbis than expected who have highly compact appellation date pairs.

An experiment for personality p consists of presenting to the decoder a list of N_p (appellation,date,compactness type) triples. The decoder then produces a table of compactness entries. For the t^{th} text and n^{th} (appellation,date,compactness type) triple, the entry is * if either the appellation or date key word have no ELSs. If both the appellation and date key words have at least one ELS, then the entry is a number representing the compactness of the result. Small numbers represent high compactness. We denote the value of the entry corresponding to the t th sampled text and n^{th} (appellation,date,compactness type) triple of the p^{th} personality by $c(p, t, n)$.

To normalize out the possibility that the compactness values for the different (appellation,date,compactness type) triples have different distributions, the table of compactness values is processed to produce a rank table. The rank table entry $r(p, t, n)$ is the number of texts that have smaller compactness for (appellation,date,compactness) triple n; i.e., the number of (n, s) pairs, where s ranges over all the texts for which $c(p, t, n)$ is not *, such that $c(p, s, n) < c(p, t, n)$. In the rank table, an entry which was a * in the raw table becomes rank T. Thus if for some triple (p, t, n), $r(p, t, n) = 0$, it means that there was no text that produced a smaller compactness than text t for (appellation,date,compactness) triple n of personality p. This happens when $c(p, t, n) = min\{c(p, t', n)|allt'\}$. If $r(p, t, n) = 10$, it means that there were 10 texts for which the decoder produced a smaller compactness value for the n^{th} (appellation,date,compactness) triple than the compactness produced by the decoder using the t^{th} text. The smallest numbered rank produced by this scheme is 0. For a non * raw entry, the largest numbered rank produced by this scheme is $T - 1$. This happens when there exists at least one ELS for the appellation key word and at

least one ELS for the date key word and there is exactly one text t for which $c(p,t,n) = max\{c(p,t',n)|allt'\}$.

For each text, we form a star team composed of these champions and determine a measure of the star team's overall compactness. We do this for star teams composed of one champion, then for two champions, ..., and finally for a star team of all possible champions. Each star team's measure of compactness is the average of the normalized rank values of the champion (appellation,date,compactness type) triples in the star team for the text. Then among all star teams for each personality and text, we select the best star team. It is the one whose average compactness, when ranked among the average compactness's of all the other texts, is the smallest. The overall star team score for a text t is the sum over all personalities of the best star team's score for text t and personality p.

Having an overall score for each text, we then compare the value associated with the Torah text with the overall value associated with the each of the monkey texts. Our final measure is the fraction of monkey texts that have better overall values than the Torah text. We test the null hypothesis at the approximately .001 significance level by seeing if this fraction is less than .001. If it is, then we reject the null hypothesis at the approximately .001 significance level.

We now repeat the description just given in a more compact mathematical notation. Let $c(p,t,1),...,c(p,t,N_p)$ be the rank compactness values associated with personality p and text t for the N_p (appellation,date,compactness type) triples. Let $i_1,\ldots,i_{N_p}$ be any permutation of $1,\ldots,N_p$ satisfying

$$c(p,t,i_1) <= c(p,t,i_2) <= ... <= c(p,t,i_{N_p})$$

The N_p star teams associated with personality p and text t are those (appellation,ate,compactness type) triples associated with indices:

$$i_1; i_1,i_2; i_1,i_2,i_3; \ldots; i_1,i_2,\ldots,i_{N_p}$$

The indices associated with the star team of M champions are $i_1,i_2,...,i_M$. The raw score associated with the star team consisting of the M champions $i_1,...,i_M$ is

$$s(p,t,M) = \frac{1}{M}\sum_{m=1}^{M} c(p,t,i_m) \tag{1}$$

The rank score associated with the star team consisting of these M champions is

$$b(p,t,M) = \#\{q|s(p,t,M) > s(p,q,M)\}$$

The rank score for the best star team for personality p and text t is

$$b^*(p,t) = min\{b(p,t,m)|m = 1,...,N_p\}$$

4 The overall score for text t is the sum of the best star team scores, taken over all personalities.

$$a(t) = \sum_{p=1}^{P} b^*(p,t)$$

The normalized rank overall score for text t is

$$g(t) = \#\{q | a(q) < a(t)\}/T$$

Under the null hypothesis and neglecting quantization, $g(1), \ldots, g(T)$ are independent and uniformly distributed random variables on the interval $[0, 1]$. If the Torah code phenomena exists in the encoding and compactness schemes being tested, $g(t^*)$, where t^* is the Torah text, should be small. Thus to test the null hypothesis at the significance level w, we see if $g(t^*) <= w$. If so, we reject the null hypothesis at the w significance level.

Furthermore, we are interested in testing, whatever the results we obtain, whether the star team compactness values for the monkey texts for each personality are uniformly distributed on the interval $[0, 1]$. We will use the Kolmogorov goodness of fit test for this purpose and test the uniformly distributed hypothesis at the .05 significance level.

If this uniformly distributed hypothesis is not rejected, then under the hypothesis of no Torah code phenomena, we can assume that the Torah text behaves just like one of the monkey texts. In this case, the probability of having its star team compactness for any personality being better than a monkey text is uniformly distributed on the interval $[0, 1]$. And we can determine under this assumption what is the probability that we would have observed k of the K personalities have a best star team compactness value smaller than half of the monkey text best star team's compactness value.

We are also interested in testing, whatever the results we obtain, whether these results for the Torah text are mainly due to a few personalities or whether they are due to a majority of the K personalities. Therefore we will determine the number k of personalities whose Torah's text star team compactness value is smaller than half the monkey text's star team compactness values. If the uniform goodness of fit test succeeds, we may assume that the probability is one half that the Torah text's star team compactness value is smaller than half the monkey text's star team compactness value. Thus, we may use the binomial distribution p value

$$p = .5^K \sum_{i=k}^{K} \frac{k!}{i!(k-i!)}$$

which is the probability that under this assumpion we would observe k or more personalities whose Torah text's star team value is smaller than half the star team compactness value of the monkey texts. If this probability is smaller than .01, we reject the null hypothesis that the observed results are due to one or only a few personalities.

To test the null hypothesis against the alternative that there are more highly compact appellation date ELS pairs, we use the best star team methodology

analyzing the appellation date pairs from all the rabbis together, rather than rabbi by rabbi.

Finally, we note that in a personal communication Professor Rips has suggested that the arithmetic mean in equation 1 be replaced by a geometric mean.

7 Conclusion

Having defined an objective gathering protocol for defining a priori related appellation and date key word pairs and a statistical technique for testing the null hypothesis that the observed compactness values of the related appellation and date ELSs are as expected by chance against the hypothesis that they have smaller values than expected by chance, we are ready to do the experiments. Experimental results will be written up as soon as they are available.

References

1. Abraham P. Bloch, *Day By Day In Jewish History*, KTAV Publishing House, Inc., New York, 1983.
2. David Halachmi, חכמי ישראל, *Hochmi Israel*, Abraham Zioni Publish House, Tel Aviv, 1957.
3. Raphael Halperin, תולדות עם ישראל אטלס עץ־חיים, *Atlas Eytz Chayim*, 10 Volumes, Tel Aviv.
4. Mordechai Margalioth, אנציקלופדיה לתולדות גדולי ישראל, *Encyclopedia Letoldot Gedoli Yisrael*, *Encyclopedia of Great Men In Israel*, 4 volumes, Yavneh Publishing House, Tel Aviv, 1986 vol 1,2, 1992 vol 3,4.
5. Abraham Stern, מליצי אש, *Melitzi Aish*, Jerusalem.
6. Doron Witztum, Eliyahu Rips, and Yoav Rosenberg, "Equidistant Letter Sequences in the Book of Genesis," *Statistical Science*, Vol. 9, No. 3, 1994, pp. 429-438.
7. Michael Drosnin, *The Bible Code*, Simon and Schuster, New York, 1997.
8. Nisan Aryeh Novick, *Fascinating Torah Prophecies Currently Unfolding*, Netzach Yisrael Publications, New York, 1997.
9. Yacov Rambsel, *Yeshua*, Frontier Research Publications, Ontario, 1997.
10. Jeffrey Satinover, *Cracking The Bible Code*, William Morrow and Company, New York, 1997.
11. *Jewish Action*, March 1998.