Fictitious Toponyms in the Responsa: Bashan’s Ruleset Revisited

Ephraim NISSAN

Abstract: Fictitious personal names and toponyms are not infrequent in legal casenotes as used for didactic purposes nowadays. There is a long tradition of fictitious names being used in the legal literature. The problem with medieval or early modern legal (here, rabbinical) responsa is that if they are used as evidence for historical purposes, as though they were chronicles, confusion may occurs. Historian Eliezer Bashan showed that this is the case, indeed, with particular reference to rabbinical responsa from the Ottoman empire where Holy Land toponyms occur. He set forth several tentative rules to decide whether a toponym is there to literally refer to the place it names, or whether, instead, the name is used fictitiously. This paper formalizes the ruleset.

Introduction

Full text information retrieval is known to focus on the text itself, with semantics being sometimes handled, but in a rudimentary way. Pragmatic factors which belong in the specialist’s common-sense of a discipline concerned with the text are also worth formalizing, whether for incorporation into automated tools, or otherwise. Disciplines relevant here include law and history, perhaps more history than law.

The locale of legal cases (or, possibly, other accounts of events) in the literature of rabbinical responsa—historian Eliezer Bashan discov-
erected (1974)—is sometimes disguised by indicating, instead, the name of one or more towns in the Land of Israel. His corpus was of Responsa from the Ottoman period, but his paper also brings examples from the Rabbinical literature which fall outside that temporal or geographic scope. If place names are sometimes fictitious, who is to tell which is what? On the last page of his paper, Bashan provided a set of heuristic decision rules. In this paper, instead, a formalism is developed, for capturing the gist of Bashan’s ruleset. I have also dealt with this problem in the much broader context of a 100-page essay, on the need to formalize qualitative reasoning on space and place for legal purposes (NISSAN 1997: Part II, Chs. 6 to 8). “Buried”, as though, in such a long work, moreover in a paper-collection that, it very soon turned out to be the case, no longer was obtainable given the exceedingly small run of the printing, the topic of this paper decidedly deserves being tackled again, on its own merits. Time is pervasive in text, and is in particular pervasive in the realm of legal texts (see, e.g., MARTINO and NISSAN, 1998). Yet, space, too, is pervasive. So is the need to identify, for that matter. In NISSAN 1995 I formalized a legal text about how to proceed when identification of individuals out of a discrete universe could not obtain, in a very special context. And then, there is the pervasive need to prove, and to reason at a meta-level about whether proof does obtain (NISSAN 2000, MARTINO and NISSAN 2001a, 2001b, 2003), and, for all of there existing formalism of proof (e.g., see Shimony and Nissan’s paper in MARTINO and NISSAN, 2001a), the debate on whether such models do yield what they promise is at times acrimonious (such is the debate among the so-called Bayesio-skeptics and the Bayesian Enthusiasts in legal theory about the scope, role, and value of forensic statistics; elsewhere I described how even Voltaire contributed to the “current” debate).

Let us focus on the corpus of responsa which Bashan was concerned with. Some toponyms can be expected to idiomatically occurs by pairs. Bashan proved that the pair of toponyms ‘Tiberias’ and ‘Sepphoris’—of two towns in Galilee that are well-represented in the Talmudic literature—were used as generic toponyms in the rabbinical legal literature from the Ottoman period. Bashan finds parallel texts, in which events are related that are set in Tiberias and Sepphoris, but can be shown to apply to the Jewish communities of Patras and Lepanto in Greece, in the context of the war of 1532. In particular, details
applying to spatial notions, e.g., distance and the time required to
cover it by walking ("half a day") do not match the geographic (and
transportation) reality of Galilee. Elsewhere, Bashan shows in his pa-
per, Tiberias and Sepphoris refer to places in, e.g., Spain. Bashan
points out that based on the toponyms in ancient legal documents,
some modern historians wrongly tried to integrate the facts reported,
in the context of the history of the named towns in Galilee.

In one text from the 16th century, Sepphoris and Jerusalem are
port cities in Algeria, even though none of the places in the Holy Land
bearing those names are on the coast: “You asked: Reuben sent a de-
posit by the intermediary of Simeon, for him to carry it to Levi in Je-
rusalem, and he went by ship, and (= but) the ship was unable to go to
Jerusalem (!!); he left the deposit in Sepphoris, with Judah”.
(quoted in BASHAN, p. 151, my trans.). In the text considered, ‘Jerusalem’
stands for a place accessible by the maritime routes, and it was only
transient circumstances (not the actual Jerusalem’s geographic posi-
tion) that prevented that particular ship at that particular time to get
there. As to the four characters’ names, ‘Reuben’, ‘Simeon’, ‘Levi’,
and ‘Judah’, they are clearly fictitious.

Bashan found it sometimes difficult to decide whether any of the
toponyms in a pair, or both of them, have to be taken to be fictitious.
Some other times, other toponyms in the context make it clear that the
country is, say, Greece. Just as the names of the principals in a narra-
tive were changed, so was the name of their cities of residence, or
anyway of the locale; yet, the name of some other places was left un-
changed: “Reuben, Simeon, and Levy are associates of Jacob, a resi-
dent of Hebron. Reuben, Simeon and Levi were residents of Tiberias,
and Reuben and Simeon used to invest their joint capital wherever
they had the opportunity of doing so: in Mystras or its neighborhood”
(quoted in BASHAN, p. 155, my trans.).

Elsewhere, all toponyms were changed; e.g., in a legal case—re-
lated by R. Samson Morpurgo of Ancona (1681–1740)—of the vicis-
situdes of a deserted wife from Sepphoris who used to work as a maid
in Shechem, and met there somebody from Gaza. “It must be assumed
that all three names refer to towns in Italy” (BASHAN, p. 157). In yet
another example, parallel texts that name ‘Beersheba’ but, corre-
spondingly, Vidin, enable Bashan (pp. 160–161) to identify the refer-
ent of the former with the latter, which is a town in Bulgaria on the
Danube, the year of the decree concerned being 1377. ‘Judaea’ and ‘Galilee’ also occur as fictitious names for places where actual events reportedly took place.

**The Formalism**

In the following, I translate, one by one, Bashan’s (p. 165) set of heuristic decision rules for deciding whether a toponym is fictitious. One rule at a time, a logic representation is proposed. The formalism proposed hereby develops no specific treatment of belief (out of the panoply of approaches from artificial intelligence).

**Rule 1:** If the responsum starts by the pair of names ‘Tiberias’ and ‘Sepphoris’, or ‘Tyre’ and ‘Sydon’, or ‘Judaea’ and ‘Galilee’, which are Talmudic idioms, then one must be suspicious about the place’s real name, and to presume that the event narrated took place other than at the locale named.

Let us sketch a possible formalism. Let $T$ denote the text considered; it is an ordered set of words:

$$T \equiv \{w_1, w_2, \ldots, w_N\} \equiv \{w_i\}$$

Let $\sigma(w_i)$ be a function retrieving a set of semantic features of word $w_i$.

Let $P(T)$ be the ordered set of all place-names in $T$.

$$P(T) \equiv \{p_j\},$$

where

$$\forall p_j \in T \land (\text{TOPONYM} \in \sigma(p_j)) \land \{w_k \mid (w_k \in T \land (\text{TOPONYM} \in \sigma(w_k)))\} = \emptyset.$$
document from a given genre, a responsum has some general schema. It has elements larger than the word unit; e.g., perhaps, a preamble.

Let us simplify, and stick to the letter of Bashan’s Rule 1. It was worded for the consumption of human researchers, not an ideal piece of software, and like all texts meant for human readers, it has to be interpreted with a grain of salt. For the purposes of our simplification when formalizing Rule 1, let us assume that we actually have to check only the first two place-names in T. As P(T) is an ordered list, we can use function car (in Lisp style) to retrieve the first element of the list. The operational rule for constructing P(T) out of T is to recursively select the first word out of (the remaining part) of T, and to retain it in the output if and only if it is a place-name. The order in P(T) reflects the order of T. Namely, P(T) is what remains of T once all such words that are not place-names are deleted. More precisely, only toponyms are retained, thus excluding, e.g., such common names as ‘country’ or ‘town’. Realistically, however, we may have to determine whether a word is a toponym based on the context, i.e., based on other words and higher syntactic units even when these are not toponyms; but let us assume that such a refinement is to be embedded in the definition of σ(wi).

Now, we can formulate the condition part of Rule 1:

\[
\text{If } \left( \text{car}(\text{P}(T)), \text{car}(\text{car}(\text{P}(T))) \right) \in S, \text{ where } S \equiv \{ \text{‘Tiberias’, ‘Sepphoris’}, \text{‘Tyre’, ‘Sydon’}, \text{‘Judaea’, ‘Galilee’} \} \text{ then …}
\]

Keep present that such a formulation does not state the justification, namely, Because \( S \subseteq R \), where R is defined as the repertoire of idiomatically paired terms from the Talmudic literature, which in turn is a major source shaping the Responsa in several respects.

In set S, for which we define this symbology:

\[
S \equiv \{ \pi_h \} \equiv \bigcup_{h=1}^{3} \pi_h,
\]

we are only considering the three pairs that Bashan stated in Rule 1. Apart from \( \pi_h \in R \), we need to represent the fact that the first word,
\( \pi^1 \equiv \text{car}(\text{car}(\pi_0)) \) in \text{LISP} symbology, and the second word, \( \pi^2 \equiv \text{car}(\text{cdr}(\pi_0)) \), in all three pairs, are Hebrew toponyms whose proper denotatum is some given place in the “Land of Israel”. The latter, for our present practical purposes, is identical with what in Western culture is known as “the Holy Land” (roughly, the country between the River Jordan and the sea), even though strictly speaking, in Halachic terminology things are not that simple (cf. the Halachic status of Akko in the gemara). Let us define the symbol \( H \equiv \text{denotatum}('\text{Holy Land}') \).

We can tentatively express the considerations made above as follows:

\[
\forall h, 1 \leq h \leq 3, \forall f, 1 \leq f \leq 2, \\
\text{TOPONYM} \in \sigma(\pi^f h) \land \text{denotatum}(\pi^f h) \subset H \land \\
\pi^f h \subset \text{Hebrew}.
\]

Having sketched the representation for the condition part of Rule 1, let us complete it now with the rest of that rule:

\[
\text{If } p^1_T, p^2_T \in S, \\
\text{where } p^1_T \in \text{car}(P(T)), \\
p^2_T \in \text{car}(\text{car}(P(T))), \\
S \equiv \{ \pi_h \} \equiv \{ \pi^1_h, \pi^2_h \} \\
\equiv \{ ('\text{Tiberias}', '\text{Sepphoris}') , \\
('\text{Tyre}', '\text{Sydon}') , \\
('\text{Judaea}', '\text{Galilee}') \},
\]

\[ \text{then } \text{Likelihood} ( (\text{referent}(p^1_T) \neq \text{denotatum}(p^1_T)) \land \\
(\text{referent}(p^2_T) \neq \text{denotatum}(p^2_T)) \leftarrow \text{high} \]

This can be augmented with the aetiology, i.e., the reasons behind the rule as stated. We omit this here. Moreover, we are not going to discuss how to formalize the treatment of likelihood degrees, for which, refer to the standard literature of artificial intelligence.

\[ \text{Rule 2: Even if the ones that appear, are two other names of towns of the Land of Israel, or, possibly, just the} \]
name of one such town, and there is no other evidence of there having been Jewish residents in those towns during the period when the *responsum* was written, then one should not use that only source as evidence that such a Jewish community was present there indeed, and one should rather assume that the name was just fictitious (e.g., Beer-Sheva).

Let us define the symbols $\mathcal{R} \equiv$ religion and $y \equiv$ Jewish. This time, just the words, i.e., the signifiers, are stated, unlike when we defined: $\mathcal{H} \equiv$ denotatum (‘Holy Land’). However, omitting a statement to the effect that it is the denotatum that is meant, is only a conventional shorthand of the representation. We only stated the denotatum explicitly to avoid confusion with the toponym. In general, you may conceive of this omission of an explicit indication, the way it is usual with representations manipulated by LISP: the data structure can be either treated as data, or evaluated, according to the operational context. Filling in the facets in a frame does not require one to point out it’s the value of the identifier that is to be accessed by the control component.

Let us sketch a formal representation of Rule 2. The rule applies to $\forall p_j, p_j \in T$, but here again, a refinement is called for. There is no way the entire text of a *responsum* can be assumed to be free of any occurrence of such a toponym that does not only denote, but also actually refer to a town that is not the relevant locale for the current events with which the *responsum* is concerned. For example, the wish may be formulaically expressed that Jerusalem will be rebuilt soon. Indeed, apart from legal cases or other current communal reports, one may expect the style to be florid here and there, interspersed with Biblical references and quotations. A cited verse could conceivably mention a place that is not relevant for the case at hand in the particular *responsum*.

Here again, this is obvious for readers of Bashan’s article, but would require explicit statement in a full-fledged formalism. Therefore, let us assume that items that appear in $P(T)$ are only such toponyms whose referents (not just the denotatum) are places of current relevance for the *responsum*, regardless of the denotatum of the toponym. If the denotatum and the referent are different, then the
toponym is fictitious, but now we are excluding even actual references to places other than specifically relevant to the responsum. Having redefined \( P(T) \equiv \{ p_j \} \) as explained, let us go on representing Rule 2:

\[
\forall p_j, p_j \in P(T),
\]

If \( \text{denotatum}(p_j) \subset \mathcal{H} \)

\( \land \) assertions\( (T) \supset \)

assertion\( (J_0) \) is resident

in \( \text{denotatum}(p_j) \)

at-time current-time\( (T) \)

where \( (\mathcal{R}, \text{of } J_0 \text{ is } y) \)

\( \land \) \( \neg \text{Known}(\exists \text{Evidence } E_0 \text{ to-the-effect-that } C_0) \)

where \( C \equiv [\exists J, (\mathcal{R}, \text{of } J \text{ is } y)] \)

— that is to say, there is no evidence other than from \( T \) that any Jewish residents were in \( p_j \) at the times of text \( T \) —

then Likelihood \( (\text{referent}(p_j) \neq \text{denotatum}(p_j)) \leftarrow \) high;

Don’t (Claim \( [E_0 \Rightarrow C_0] \))

Note that the then-part is an action-part. The one instantiated here comprises an assignment statement, along with a negative prescription. In this representation, the scope of Known is unspecified, and we are not going to delve, either, into a discussion of whether to go the modal or the ontological way, i.e., whether to treat epistemic states by means of operators or, instead, as predicates. One remark is nevertheless called for, concerning the scope of Known. For practical purposes, the delimitation of the knowing agent includes the inquiring agent at hand, i.e., the historian resorting to the ruleset. Arguably, some form of consensus is also called for: whether it is knowledge from the published literature, or the historian’s own working records or mental states; and whether Known applies to an uncontroversial proposition, or merely to a belief that is entertained as a working hypothesis.

If we are to be punctilious about getting a neat theoretical framework, we may muse as well about how to represent a common belief. The concerns of some current theoretic research are clearly beyond the practical needs for the case at hand. “The notion of common belief is infinite in nature. It means that everybody believes some proposition

---

or event, everybody believes that everybody believes it, etc. To what extent can this notion be grasped by an epistemic logic that admits only finite conjunctions?”, asks Aviad Heifetz at the start of a paper (1996), “Common belief in monotonic epistemic logic”, in which he tries to solve the problem stated. He concludes “that finitary epistemic logic cannot capture the full contents of the notion of common belief” (p. 122). Now, let us turn to the next rule in the ruleset, namely, Bashan’s Rule 3 (my trans.):

**Rule 3:** If two names of towns in the Land of Israel are mentioned, and of one of these towns it is clearly known that no Jews were residing there (e.g., Ascalon), and the respondent nevertheless wrote that they did; then, in the responsum considered, fictitiousness should be assumed also for the second place, even though Jews are known to have been residing there (e.g., Gaza).

A formalized representation of Rule 3 is given below:

\[ \forall p_u, \forall p_v, p_u \in P(T), p_v \in P(T), \]
\[ \text{If} \]
\[ \text{denotatum}(p_u) \subseteq H \]
\[ \wedge \text{denotatum}(p_v) \subseteq H \]
\[ \wedge \text{Known}( C(p_v) \wedge \neg C(p_u) ) \]

where \( C(\pi) \equiv [ \exists J, ( \mathcal{R} \text{ of } J \text{ is } y) \]
\[ \wedge ( J \text{ is resident in } \text{denotatum}(\pi) \]
\[ \text{at-time current-time}(T) ) ] \]

then Likelihood (referent(p_u in T) \neq \text{denotatum}(p_u)) \wedge
\[ (\text{referent}(p_v \text{ in } T) \neq \text{denotatum}(p_v)) \] ← high

The next rule in Bashan’s ruleset is Rule 4, which I translate as follows:

**Rule 4:** If one is faced with the name for a given town in the Land of Israel, of which it is known that there used to be there, indeed, Jews in residence, but the details given along are not realistic (e.g., four separate Jewish communities in Shechem [i.e., Nablus]), then one
must assume that the reference is not at all to that town.

\[ \forall p_w, p_w \in P(T) \]
\[ \text{If } \text{denotatum}(p_w) \subseteq \mathcal{H} \quad \land \quad \text{Known}(C(p_w)) \]
\[ \text{where} \quad \text{[Def. of } C(\pi) \text{ is in Rule 3]} \]
\[ \land \exists \epsilon, \epsilon \subseteq \text{assertions(on } p_w \text{ in } T) \]
\[ \land \text{Known}(\Phi) \]
\[ \land \text{Known}(\neg \text{Defeasible(Known}(\Phi))) \]
\[ \land [\epsilon \Rightarrow \neg \Phi] \]
\[ \text{then } ( \text{referent}(p_w \text{ in } T) \neq \text{denotatum}(p_w) ) \]

The latter heuristic rule may be expressed in some logic approach among those whose stance is the usual one, of eradicating contradiction (e.g., as in ALCHOURRÓN \textit{et al.} 1985), rather than allowing to retain inconsistency (cf. BLAIR and SUBRAHMANIAN 1989, GHOSH 1994, PEQUEÑO and BUCHSBAUM 1991). This consideration applies at least to the extent that it is the conclusion-part of Rule 4 that is concerned. A more general formal framework for historical reasoning should, however, specifically address the need for handling paraconsistency.

**Rule 5:** If it is known that such details that are reported as having taken place in a certain locale in the Land of Israel, are identical with historical details that are known from other historical sources as having taken place abroad, and these events involve, for example, a character who is known to have been active in a certain community abroad (e.g., Rabbi Y. Pyrmwn in Patras), then, clearly, in the case at hand the toponym or toponyms is or are fictitious.

\[ \exists \delta, \kappa, p_D, p_A, \]
\[ \land \kappa \text{is-a person} \]
\[ \land \delta \subseteq \text{assertions(in } T \text{ on } \kappa) \]
\[ \land \lambda \equiv \{ \forall e \mid [e \text{is-a Evidence(on } \kappa)] \land 
\[ [(\neg e \text{ in } T)] \}
\[ \land \lambda \neq \emptyset \]
\[ \text{Known} \left( \text{referent} \left( \kappa \text{ in } T \right) = \text{referent} \left( \kappa \text{ in } \lambda \right) \right) \]
\[ \text{Known} \left( \neg \text{Defeasible} \left( \text{Known} \left( \text{referent} \left( \kappa \text{ in } T \right) = \text{referent} \left( \kappa \text{ in } \lambda \right) \right) \right) \right) \]
\[ \exists N \in N \text{ is-a Narrative} \left[ \text{Protagonist: } \kappa, \right. \]
\[ \left. \text{Events: } \text{Events} \left( \text{of } N \right), \right. \]
\[ \left. \text{Locale: } \text{Locale} \left( \text{of } N \right) \right] \]
\[ \land N_1 \text{ is-a } N \land N_1 \subseteq \delta \]
\[ \land \text{Locale} \left( \text{of } N_1 \right) = p_D \]
\[ \land N_2 \text{ is-a } N \land N_2 \subseteq \lambda \]
\[ \land \text{Locale} \left( \text{of } N_2 \right) = p_A \land p_A \not\subseteq H \]
\[ \land \text{Events} \left( \text{of } N_1 \right) = \text{Events} \left( \text{of } N_2 \right) \]
\[ \land \exists p, \left[ \left( p \in P(\lambda) \right) \land \left( p \subseteq H \right) \right] \]

then \[ p_A = \text{referent} \left(p_D \right) \neq \text{denotatum} \left(p_D \right) \]

Bashan’s Rule 6 involves privacy considerations as being the motivation for disguising the toponym:

**Rule 6:** If it is reprehensible or ordinary but private sexual matters that allegedly took place in some town of the Land of Israel, then one must suspect that there was an intention to hide the actual names of the locale.

Cuppens and Demolombe (1996) have defined a deontic logic for reasoning about confidentiality. Their application is database security, and their formalism is a device for expressing security policies for guaranteeing confidential access to the database.

Instead, I propose this *ad hoc* representation for Bashan’s Rule 6. Let us indicate deontic necessity by \( N_D \). Then, we can tentatively formalize Rule 6 as follows, even though more narrative structure is perhaps called for, along with a more accurate description of epistemic states:

**If** \[ \exists \mu, \nu, p_m, \]
\[ \mu \subseteq \text{assertions} \left( \text{in } T \right) \]
∧ p_m ∈ P(µ)
∧ denotatum(p_m) ⊂ H
∧ ν ⊂ µ
∧ themes(ν) ∋ Intimacy
∧ N_D(¬ν)

then Likelihood (referent(p_m) ≠ denotatum(p_m) ) ← high

A formal representation is also called for, in order to represent the etiology of this. A goal-driven structure would be suitable: suppressing the actual toponym is an available plan in order to prevent the characters in the narrative from being traceable. This, in turn, is motivated by the wish to avoid potential harm to the reputation of families, whether correctly identified or not.

The last rule stated by Bashan (p. 165) is the following Rule 7. This translation reflects the fact it was not originally formulated as being explicitly constituted of a condition part and an action or conclusion part. It is nevertheless amenable to such a structure.

Rule 7: There may occur linguistic indicators to the effect that the respondent only meant to use the toponym out of convenience in a hypothetical (e.g., “If in Tiberias there are three Holy Congregations”).

Information Retrieval and Concluding Remarks

In this paper, I formalized a set of rules that was proposed by Bashan, a historian, in order to tentatively determine whether in legal casenotes or reports of episodes, the occurrence of names of towns in the Holy Land actually refers to those towns, or was used instead fictitiously, as a cover for the actual toponyms, the identities of the actual places being hidden for some reason or just omitted because authors did not deemed it interesting.

The seven rules analyzed are merely heuristic rules. Bashan concludes (p. 165, my trans.):

Yet, not always things are clearly cut, and there are as well such borderline cases of which it is difficult to decide whether the name is the real one. Anyway, a close examination supported by
experience and subtle discernment should put the scholar in a position to tell which is what: whether the toponym is just fictitious, or it is there for real.

Needless to say, several phenomena in common sense or in historical inquiry methodology are not accounted for, in the present paper’s formalism. We didn’t delve into theoretical issues of the representation, either, even though more thorough a discussion is warranted, in an artificial intelligence perspective, concerning the representation of knowledge of a proposition, autoepistemic assumptions about its not being defeasible, and the comparability of the likelihood of different kinds of propositions.

I would like to note in addition that a different kind of ambiguity between names for towns is instantiated by \(<q_wșt'>\). This is how Qușța, the name of a legendary City of Truth (whose inhabitants only speak truthfully), is written in the Babylonian Talmud, Tractate Sanhedrin, folio 97, side a, its quite transparent etymology being the Aramaic noun for ‘truth’, which is qușța (written \(<q_wșt'>\)) or qșot (written \(<qṣwț'>\)). The text of the Talmud is in Aramaic and Hebrew.

As could be expected, the Aramaic common name for ‘truth’ co-occurs with the toponym in Sanhedrin, 97 a: “formerly I thought there was no truth in the world”, versus “he came to a place called Qușța”.

Location is also involved in the standard phrase (extant in both Aramaic and Hebrew) ‘the world of truth’, which means ‘afterlife’. On the legend of the City of Truth, see WEISS 1985, p. 125, fn. 23, and the literature cited there.

However, ‘Kosta’ is also a shortened form of both the personal name ‘Constantine’ (and as such it occurs in the Modern Greek homomasticon), and the name of Constantinople. In Hebrew texts throughout Byzantine and then Ottoman history, more than one form of the name for Constantinople occurs, but anyway in the 20th century in formal writing in Hebrew, any reference to Constantinople (at least in a late Ottoman historical context) rather than Istanbul (the capital of republican Turkey) is by the name Qoșța, written \(<q_wșt'>\). In this example, the names for the City of Truth and of Constantinople are unrelated, and merely a coincidence in the spelling occurs. The textual corpora in which either city is mentioned are disjoint, and belong to different historical ages. Moreover, the Talmudic occurrence, a hapax in context, is unmistakable, and moreover the difference between the
toponym and the common name for ‘truth’ is explicitly indicated in the text. It’s not like the name for an extant city is being used fictitiously.

Schweighofer (1999) is concerned with information retrieval from legal texts; the following explains the difference between homonymy and polysemy, and points out the problem that polysemy poses for information retrieval from texts whose genre is legal:

Homonyms are words that match in spelling but have different meanings […]. Polysems are expressions with initially the same importance, however by transfer, analogy, historical or regional development have developed different meanings or are employed as general terms in different contexts. For this problem the Boolean search in inverted files offers no solution since words with the same spelling are taken as homonyms.

In jurisprudence, polysems are very frequent but represented with other characteristics. Many legal terms are employed in a different context than in natural language. The exact importance of a legal term can only be understood in its context. This phenomenon is considerably more frequent than the similar homonymy and substantially affects the quality of results in legal database queries.

Homonymy and polysemy depress precision because such words have more than one meaning. Sense resolution will be an important component in future retrieval systems. The disambiguation techniques are mainly based on the idea that a set of words occurring together in context determine appropriate connotations. This approach can be used to compute various senses of descriptors […] or determine appropriate senses for word sets despite each individual word being multiply ambiguous.

That <qwšt’> may refer to Constantinople or, instead, to the City of Truth is a case of homonymy. That in a 16th-century document, the Hebrew names for Sepphoris and Jerusalem are used to refer to two port cities in Algeria is not even a case of polysemy. Rather, it’s a case of non-literal use.

In this paper, I have described a formal device for word-sense disambiguation in a restricted semantic category: a set of toponyms, subsets of which (especially doubletons: pairs of toponyms) may be used fictitiously, instead of by bona fide reference, in a textual corpus which is itself delimited by genre, and geohistorical context of origination. The disambiguation device heavily relies on pragmatic factors,
as well as on encyclopaedic knowledge and on the episodic memory and knowledge structure which obtain from a particular document. The insights yielded by this discussion stem from a peculiar subject, yet they hopefully contribute to an evaluation of what, in general, can be expected to be involved in word-sense disambiguation while processing (mentally or mechanically) a text shaped by generic conventions. The interest of the formalism described in this paper is theoretical, yet, being applied as it does to rabbinical responsa, a domain which can boast of signal achievements in information retrieval thanks to the work of Prof. Yaakov Choueka, who in the 1980s supervised my doctoral project (which in turn was on a computational model of word-formation: see, e.g., NISSAN 1987, 1988, 1991a, 1991b, 1992a, 1992b, 1999b, 1999c, 1999d), the ambition is not out of place that this study may also be considered as a probe into the potential of semantic processing of the same corpora. In another project, I have dealt with terminological doubletons, in the form of word-formation in response: more in particular, what I call misantonyms (NISSAN 1999a). In the present paper, we have seen that often fictitiously used toponyms also come in doubletons.

The fictitious use of toponyms in rabbinical responsa from Ottoman lands, as shown by Bashan, can be read correctly once one is aware of the convention in the corpus, or indeed more generally in the genre. The related genre of modern legal casenotes resorts to invented toponyms (e.g., during most of the 20th century in Italy the broad public has been finding such casenotes in the country’s leading crosswords magazine, La Settimana Enigmistica).

In my COLUMBUS model (NISSAN 2001b, cf. NISSAN 2001c) I have shown how the poetic conventions of a genre shape both narration and explication, and how this can be formalized in a notation amenable to partitioned semantic networks. In my discussion of the JAMA legal narrative and model, I have shown how a repertoire of narratives from collective memory or from the literary canon of a modern culture at a given historical moment may shape public perceptions of a criminal event and its handling by the state, as conveyed by the media (GEIGER et al. 2001, NISSAN 2001a).

The following is taken from Nissan (1997: Part II), and point out the interplay of phonemic stress, morphology, and lexical semantics when a word in a text is ambiguous:
Problems with ambiguity may involve toponyms in several ways. Here is an example which involves words not found in either a gazetteer or a dictionary, but such that can be formed morphologically, by either inflection or derivation. Moreover, orthographic practices are involved: in Italian, the tonic stress is usually not written, other than in such a word (other than a monosyllable) that ends by a vowel, if that vowel is stressed. From context, a fluent reader of Italian can decide whether, say, aguzzino is to be stressed on the i (thus, aguzzino), being understood as a noun (for ‘torturer’), or to be rather understood as a verbal form, stressed on the u (thus, aguzzino), and meaning “that they sharpen”.

Let us assume that, in an Italian text, the word arieggino occurs. Nobody fluent with the language would fail to read this word with the stress on the e. Indeed, this word would be interpreted as being regularly formed, by inflection, as the 3rd person plural of the subjunctive in the present tense, from verb arieggiare, which is extant and listed in the dictionaries. The sense would be understood to have to be picked out of “[that] they air/ventilate”, or “[that] they resemble/look like” (if used transitively); or, if the verb occurs with the preposition a, the sense would be either “[that] they resemble/look like”, or “[that] they imitate/put on the airs [of]”.

However, French Ariège is the name of a river as well as of the department, bordering with Spain, named after that river. Hypothetically, if the Italian text was to mention by name il dipartimento dell’Ariège “[the department of Ariège]”, and then we were to find there an occurrence of arieggino, it would be legitimate to wonder whether the word should be read with the stress on the i, by taking the word to be an adjective for the region, or a name for the inhabitants.

True, the word is not part of the extant Italian lexicon, but it is part of the potential lexicon, as it is transparently formed, by morphological derivation, out of a proper name that is clearly relevant in the context of the document. It’s unnecessary to know that the equivalent French word is ariégeois (for the inhabitants, or, as an adjective, as in les Pyrénées ariégeois). The derivation of the Italian arieggino (i.e., arieggino) from Ariège would be supported, by analogy, by the extant viareggino from the toponym Viareggio.

Information retrieval (IR) may be dumb and exhaustive (which is precisely why information retrieval on the responsa has been so useful for linguist Kaddari (1991–1994) to describe Hebrew syntax throughout its historical stages). Concordance-wise, all contexts containing an
occurrence of a string, or of an inflected form of a lemmatized item, or of a lexical derivative of root, will be retrieved. Or, then in contrast, devices for pruning may be introduced which attempt to guess some semantic feature, yet may result in omissions undesirable for given purposes. The latter shortcoming is amenable to the tentative nature of relevance criteria, and to Swanson’s (1988) Postulates of Impotence, which “apply only to the problem of subject-oriented information retrieval”:

PI 4: It is never possible to verify whether all documents relevant to any request have been found […].

PI 5: Machines cannot recognize meaning and so cannot duplicate what human judgment in principle can bring to the process of indexing and classifying documents […].

PI 8: You can have either subtle relevance judgments or highly effective mechanized procedures, but not both […].

Deep semantic analysis is feasible, for multi-paragraph text, in computational linguistics; this is not the case when large amounts of texts are involved. At the very least, techniques of automated story categorization are important, for hopes of future automated processing making use of the formalized version of Bashan’s ruleset. At the meet of information retrieval and natural-language processing, ambiguity may affect story categorization, not just the lexicon (or, for that matter, how syntax is parsed). Hayes et al. (1988) reported about a tool that, while “not performing a complete semantic or syntactic analysis of the input stories”—a stream of news stories—was sophisticated enough to handle such lexicon that its superordinate semantic concepts would make for misleading predictions about the category in which the story belongs.

We were prepared for sports stories that looked like metals stories (“… captured the gold medal at the summer Olympics…”) or like war/disorder stories (“…the battle on center court at Wimble-
don…”). A more difficult challenge was posed by words and phrases that were good predictors of a particular topic but occurred randomly across all story types, sometimes with the same meaning, sometimes not. […] Metaphorical language was also a problem— not use of fixed phrases (we had no trouble failing to assign the category metals to a story that contained the phrase like a lead bal-
loon)—but rather creative metaphorical language. So, a story about
a series of battles in the continuing disposable diaper war between Proctor and Gamble and its competitor was assigned to the disorders category.

Whereas semantic processing can boast of advances over the last two decades, as far as natural-language processing is concerned with short text, of at most several paragraphs, “Hic sunt leones” is still an apt descriptor on the map of desiderata, when it comes to labelling the goal of reliable semantic processing of large textual corpora. The following is quoted from an overview section in Schweighofer (1999):

So far, the experiments with word sense disambiguation techniques have been disappointing. Voorhees (1993) uses the word senses in the lexicon WorldNet for disambiguation. The improvement of the retrieval is modest. The sense disambiguation is very difficult in short query statements. Missing correct matches because of incorrect sense resolution have a deleterious effect on retrieval performance. Sanderson (1994) uses pseudo-words for disambiguation. [Cf. Sanderson (2000).] The results on retrieval performance are similar to those of Voorhees. An implementation can only be recommended if the disambiguator is able to resolve word senses to a high degree of accuracy. The NetSerf application of (CHAKRavarthy, Haase 1995) is a programme for searching in information archives on the Internet with semantic knowledge representations and with a disambiguator using the lexicon WorldNet. Semantic knowledge representations of archives lead to a remarkable increase of performance. The disambiguated version performs slightly worse than the undisambiguated version. Disambiguation with the help of a lexicon or a text corpus provides a strong impetus to the goal of multilingual information retrieval (see the proceedings of SIGIR’96 and SIGIR’97, especially (Davis, Ogden 1997, Hull, Grefenstette 1996)). The situation is different if the technique of word sense disambiguation is used for automatic indexing. Quite promising results were achieved with disambiguation (Stairmand 1997) and in the KONTERM project.

Schweighofer (1999) is about KONTERM, a tool which resorts to neural networks for the purposes of clustering, and whose application is to large legal documents. Aspects of KONTERM include word-sense disambiguation as well as “document description and the automatic generation of hypertext links”. Other advanced IR tools exist, which use rules and templates, and whose performances are some-
times in some relation of being partly amenable to tasks of automatic summarization (on the latter, see MANI 2001). On IR tools, I am quoting again from Schweighofer (1999):

In RUBRIC (TONG et al. 1987, TONG et al. 1989), the conceptual description of a legal domain is transformed into rules for recognising templates in documents. Probabilities can be expressed as evidence rules in order to achieve several conclusions depending on the degree of belief. […] The templates are formalised with Boolean logic. […] The belief functions allow the interpretation of Boolean operators as fuzzy-set operators. […] SPIRE […] is a hybrid case-based reasoning and IR system locating passages with interesting factors in fulltext judgements (Rissland, DANIELS 1995, DANIELS, RISSLAND 1997). The knowledge base is organised like in the case-based reasoning system HYPO. SPIRE computes the similarity between the problem and the cases of the knowledge base. Best cases are used to produce automatically a query in the INQUERY IR system. SPIRE can rank the most important passages (information extraction system). Within the relevance feedback, the user can mark relevant passages or improve the given representation of factors. Feature vectors are used to extract relevant factors from fulltext decisions for the case-based reasoning system HYPO (BRUNINGHAUS, ASHLEY 1997). For each relevant factor, a weighted feature vector is defined. A learning phase determines threshold values for classification. The attached factors are used for the computation of case similarity. […] [I]n the SALOMON project (MOENS et al. 1997 [and 1999; UYTTENDAELE et al. 1998]), Belgian criminal cases are summarised automatically and presented by a case profile. Case category, case structure and irrelevant text units are identified based on a knowledge base represented as a text grammar. Thematically important text units and key terms are selected by shallow techniques. These paragraphs are represented as weighted vectors and clustered. The system points the user effectively towards relevant texts.

About the Responsa Project, Schweighofer (1999) only cites Choueka (1980), by Prof. Yaakov Choueka, director of the Responsa Project in 1974–1986, and director of the Institute for Information Retrieval and Computational Linguistics of Bar-Ilan University. In Schweighofer (1999), there is no reference to later work from the same team (e.g., ATTAR et al. 1983, KIRSH 1998) and to the breakthrough that occurred since, when Choueka and his his collaborators
were finally able to crack the nut of disambiguating words in the short context, so that the vowels could be put back in Hebrew text, whose usual spelling conventions mostly omit them. Hebrew text can be considered to be a kind of very ambiguous text. If people manage to read it at all, it’s because in Semitic languages, word-forms are not very free, and morphological, syntactic, and semantic levels of processing cooperate towards the text being read. That with NAKDAN and NAKDAN TEXT (CHOUeka and NE’EMAN 1995), the proper vowels (represented by traditional diacritic signs) can be plugged back into the text, is a signal victory on the path to processing the semantics of the text. For a vowelless word in isolation, Choueka’s NAKDAN program generates all of the possible vowelled words, but only those that are acceptable in the language, at a 100% success rate. Vowelling in context is much more complex, as it involves disambiguation; the NAKDAN TEXT program reportedly had a success rate of 95% at vowelling vowelless textual passages. Those tools are themselves part of Choueka’s RAV MILIM project at the Center for Educational Technology in Ramat-Aviv, a project that among the other things resulted in a dictionary based on lexicographical criteria similar to the Collins COBUILD’s. I have provided an annotated bibliography of tools for Hebrew in Lancashire and Nissan (1991).

It’s against this backdrop that the formalism of the present paper was developed. It certainly didn’t originate in a vacuum, and likewise it’s my hope that it will not be merely a voice calling out in the desert. The device is a building block along with other techniques. Taken individually, they cannot go far, but eventually an overarching architecture should make good use of them. Introducing Swanson (1988), reprinted in the “Envoi” section of their readings volume, Spark Jones and Willett (1997, p. 553) stated that for all of “Swanson’s downbeat assessment of the limits of the feasible” (in his Postulates of Impotence about information retrieval), a positive interpretation can be afforded. Swanson’s own paper concluded with these words: “The potential for improving human interaction with recorded knowledge is immediate, but there seem to be few signs of interest in so practical a goal. Waiting for Godot, we fail to grasp what is now in reach”.
References


CUPPENS (Frédéric) and DEMOLOMBE (Robert): 1996. “A Deontic Logic for Reasoning about Confidentiality”. In BROWN (Mark A.) and CARMO (José) eds: *Deontic Logic, Agency and Norma-


HULL, (David A.) and GREFENSTETTE (Gregory): 1996. “Querying Across Languages: A Dictionary-Based Approach to Multilin-
gual Information Retrieval”. In *Proceedings of SIGIR ’96*, pp. 49–57.


NISSAN (Ephraim): 1987. “ONOMATURGE: An Expert System for Word-Formation and Morpho-Semantic Clarity Evaluation” (2 parts). In: CZAP (Hans) and GALINSKI (Christian) eds: *Terminol-

NISSAN (Ephraim): 1988. ONOMATURGE: An Expert System in Word-Formation (in English, 3 vols., ca. 600 pp.). PhD Dissertation in Computer Science (Beer-Sheva, Israel: Ben-Gurion University of the Negev, [IPA Award in Computer Science].


NISSAN (Ephraim): 1999b. “Registers of Use, and Ergolectal Versus Literary Niches for Neologizing Creativity. What Do the Makers of Technical Terminology Stand to Learn From Such Contrastive Analysis?”. In Proceedings of the EAFT “Conference on...


SPARK JONES (Karen) and WILLET (Peter) eds: 1997. *Readings in Information Retrieval* (San Francisco: Morgan Kaufmann).


