Scientometrics
DOI 10.1007/s11192-011-0349-2

Interpreting maps of science using citation context sentiments: a preliminary investigation

Henry Small

Received: 5 October 2010

Abstract It is proposed that citation contexts, the text surrounding references in scientific papers, be analyzed in terms of an expanded notion of sentiment, defined to include attitudes and dispositions toward the cited work. Maps of science at both the specialty and global levels are used as the basis of this analysis. Citation context samples are taken at these levels and contrasted for the appearance of cue word sets, analyzed with the aid of methods from corpus linguistics. Sentiments are shown to vary within a specialty and can be understood in terms of cognitive and social factors. Within-specialty and between-specialty co-citations are contrasted and in some cases suggest a correlation of sentiment with structural location. For example, the sentiment of "uncertainty" is important in interdisciplinary co-citation links, while "utility" is more prevalent within the specialty. Suggestions are made for linking sentiments to technical terms, and for developing sentiment "baselines" for all of science.

Keywords Citation contexts · Maps of science · Sentiment analysis · Corpus linguistics

Introduction

The analysis of sentiments in linguistics usually refers to the detection of some form of affect or emotion in language, for example, expressions of approval or disapproval (Khan et al. 2009). In scientific writing and particularly journal articles, however, sentiments are often difficult to detect (Verlic et al. 2008), and can be suppressed or equivocal in nature (Harwood 2009). For example, criticism of others is often hedged (Hyland 1998). One way of dealing with this limitation is to expand the notion of sentiment to include a much broader range of academic expression and more subtle ways of using and describing knowledge, including many of the rhetorical devices that have been identified in scientific

H. Small
Thomson Reuters, 3501 Market Street, Philadelphia, USA

H. Small (✉)
SciTech Strategies, Inc, Berwyn, USA
e-mail: hsmall@mapofscience.com

Published online: 10 February 2011

text (Swales 1990). Whether this broader approach can still be called sentiment analysis, or what it should be called, is open for discussion.

This paper focuses specifically on sentiments in the broad sense of attitudes and dispositions toward cited work as expressed in citation contexts, that is, the text surrounding references in scientific papers. Teufel (2010) has done important recent work on the machine recognition and automatic classification of references in scientific papers. Her work is in the framework of text understanding and the rhetorical structure of scientific papers in general, and involves what she calls "knowledge claim attribution" coupled with "citation function classification" in which automatic methods are used to classify the reasons for citing, drawing on earlier citation classification ideas.

Expanding on this work, if citation contexts are seen as commentaries on earlier scientific literature (Fig. 1), then the totality of such commentaries constitutes a collective view of prior cited literature. Aggregating contexts from current citing papers then provides a synopsis of that knowledge and attitudes toward it. Thus, citation contexts offer a means of summarizing knowledge in different parts of the scientific landscape (Small 1986).

Support for the lexical analysis of citation contexts can be found in the later writings of Kuhn (2000, pp 103–104). In a lecture given in 1990 he introduces the idea of a lexical structure to represent a scientific specialty or paradigm. Kuhn said: "Cognitive evolution depends … upon the exchange, through discourse, of statements within a community." He defines the lexicon as "… the unit which embodies the shared conceptual or taxonomic structure that holds the community together…". Of course, citation is an example par excellence of the exchange of statements within a community.

We can also exploit the co-usage of references shown in Fig. 1, to group documents into clusters or, "communities", and create a geometry of these objects in an abstract space. Co-usage patterns may also be mined for sentiments as well as content. Our discussion is restricted to maps created by co-citation, but sentiment analysis might also be applicable to other linkage forms such as co-word analysis or other citation based linkage methods such as direct citation, as Hargens (2000) has done with introductory paragraphs. The criterion

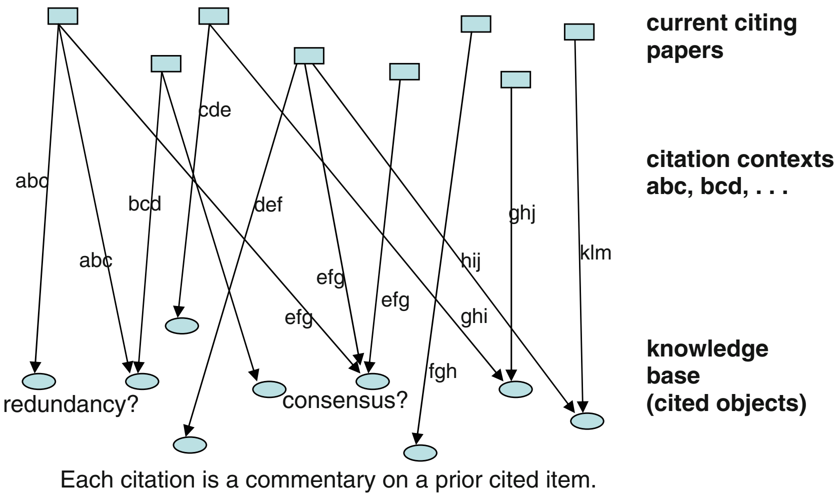

Fig. 1 Citation as commentary on existing knowledge

of whether this "linguistics of usage and co-usage" is applicable is whether the objects of discussion are embedded in a running text.

Science maps have now been created by many researchers, and the methodology continues to be improved upon (Börner et al. 2003; Chen et al. 2010; Klavans and Boyack 2009). However, maps have lacked a qualitative dimension. Technical content and subject matter have been depicted by labeling maps, but we do not know how this content has been used by the citing authors or how they regard it. Of course we want to study this qualitative dimension in a quantitative way if possible.

By combining science mapping with the linguistic analysis of citation contexts, we deepen our understanding of the structure in maps and also shed light on the cognitive and social processes that underlie scientific work. For example, we can ask how sentiments vary over the structure of a specialty from document to document, or differ from one specialty to another? What sentiments are distinctive to specialties compared to interdisciplinary science, or among different interdisciplinary links? Our hypothesis is that different sentiments are correlated with different structural locations.

This is a comparative exercise because we need to demonstrate that sentiments differ in a significant way across different portions of a map. We can think of these comparisons as either generic or subject specific. An example of a generic comparison is whether within-specialty citations differ in sentiment from between-specialty citations? Another generic comparison, not discussed in this paper, is early versus late citation (Cozzens 1985). Most comparisons are, however, subject specific, and contrast different scientific content, for example, sentiments for a specific document in the context of a specialty, or comparisons between specialties in different fields.

Identifying citation sentiments is reminiscent of earlier efforts to classify the reasons for citation (Moravcsik and Murugesan 1975; Chubin and Moitra 1975). In this early work the motive was to critique citation evaluation, but in mapping, the aim is to characterize the linked knowledge claims either from a cognitive or social perspective. Linguists and computer scientists such as Teufel have begun to explore these issues, and their work can be highly sophisticated making use of machine learning to assign text to a given set of categories based on training sets (Argamon et al. 2008). The work reported here is more exploratory and inductive, because we first need to determine what sentiments are relevant to citation contexts in different structural locations. Rather than working with a fixed set of categories, we allow the data to suggest new types of sentiments that flow from the language itself.

Method

The approach is to use citation and co-citation data to define a map, and then analyze samples of citing and co-citing contexts to find sentiments from different portions of the map and look for significant differences. Thus, the focus is on the collective behavior of authors citing and co-citing earlier literature. Access to multiple usages of the same documents, allows us to look for common patterns. For example, having multiple citation contexts for a document, make it possible to determine the degree to which citing authors share a common view of the document (Small 1978).

The first problem is how to delineate the contexts in the citing text. In this analysis contexts have been manually identified but generally they do not exceed one to three sentences around the footnote number or author-year indicator, depending on the reference style of the journal. Care was taken not to attribute text that does not apply to a given reference. References to other papers can be used to some extent to demarcate context boundaries, as well as other rhetorical cues that indicate continuation, such as “also” or “in addition”. In our work contexts consisted on average of 1.6 sentences around the point of reference.

Citation contexts provide two kinds of vocabulary: technical words and words that signal sentiments or characterizations of prior knowledge. To analyze this vocabulary we need to differentiate between words that appear frequently in one set of contexts from those that appear frequently in a contrasting corpus. Such a capability is provided by the techniques of corpus linguistics, and available in software such as Wordsmith Tools (Scott 2010). This software allows us to identify so-called “keywords” that occur more often in one corpus than another using a log likelihood ratio statistic.

We can contrast the word oriented analysis used here, where all words in a set of contexts are rolled together into a single text or corpus, with an individual context oriented approach where counts are made of the number of contexts containing one or more relevant words. The word approach is easier to use for exploration of potential candidate sentiments because keyword lists can be easily generated and scanned for sentiment bearing terms. However, the limitation of the word counting approach is that a few contexts may contribute disproportionately to the sentiment. The problem of the uneven distribution of target words in the corpus has been discussed in the corpus linguistics literature (Paquot and Bestgen 2009). Ideally counts should be made both at the word and the context level. Also, it seems clear that sentiments are not mutually exclusive, and a given context can express multiple sentiments.

Citation contexts and sentiments internal to a specialty

First we examine sentiments internal to a scientific specialty, that is, paradigmatic science, as represented by a document co-citation map on organic thin-film transistors (Fig. 2) consisting of 26 highly cited papers from 1999 to 2004 (Small and Upham 2009). There were 228 citing papers during this period for the 26 core documents, and 87 of these co-cited two or more core items in the cluster. Full texts for 81 of the co-citing papers were obtained, and these gave 304 citation contexts, two contexts for each co-citation. Only the strongest links per node are shown on the map although the nodes are 85% connected.

Of the 304 contexts, only 234 were unique which means that sometimes two or more core papers are cited within and share the same context, and thus citing authors treat them as equivalent. This is the well-known redundancy effect that Moravcsik observed in 1975. The overall redundancy rate is 23% which may be an important characteristic of within-specialty co-citations. A related issue is how close together co-citations occur in texts. 35 papers co-citing only two core documents were examined to determine whether the items were cited in the same sentence (which includes redundant cases), in the same paragraph, same section, or different sections of the paper. For this sample, 89% were same-section co-citations or closer. It might be argued that the further citations are from one another, the weaker the co-citation, but no attempt was made to quantify this effect.

Across the 234 citation contexts, there are a total of about 7,800 running words and 1,400 unique words. The top of the list is of course dominated by what are called function or stop words. The word “organic” is the top technical term occurring 98 times (Table 1). The top 10 technical words, without applying any stemming or lemmatization, are listed on the left. Descending the list we see the gradual emergence of a technical vocabulary for the specialty. At the right are listed a selection of other terms, including some stop words that

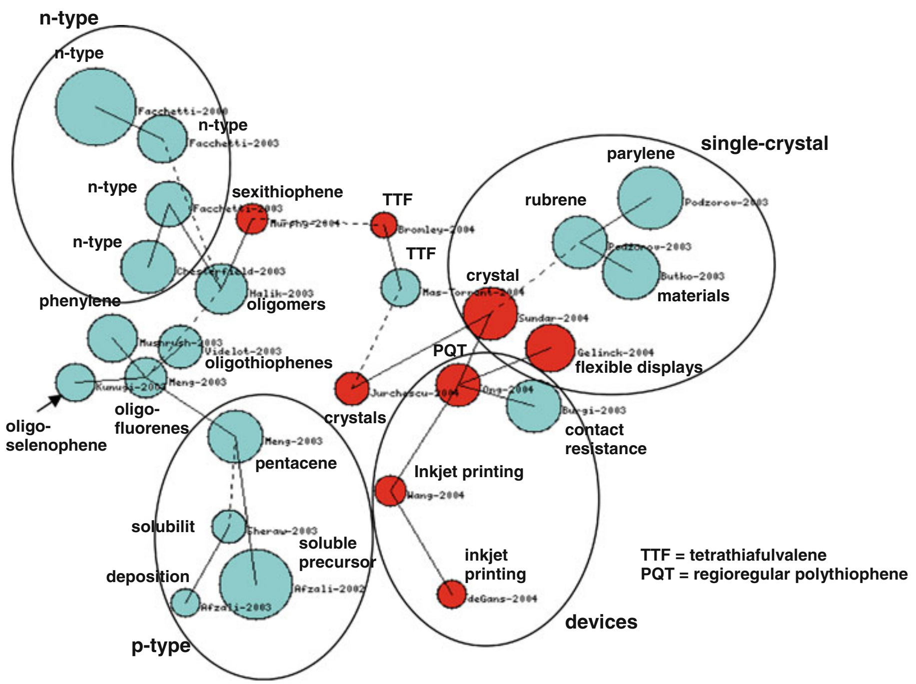

Fig. 2 Organic thin-film transistor cluster labeled by key technical terms

Table 1 High frequency words for organic thin-film transistor citation contexts

Technical word	Frequency	EGAP words	Frequency
Organic	98	As	78
Mobility	67	With	77
Materials	63	Recently	38
Pentacene	58	Reported	38
Single	53	Based	28
High	52	Using	23
Field	46	Recent	19
Type	46	All	18
Mobilities	45	Most	18
Transistors	44	Used	18

might, however, be relevant to sentiment. Linguists sometimes call these English for general academic purpose (EGAP) words (Paquot and Bestgen 2009).

With sets of citation contexts for each highly cited paper, it is possible to identify the most significant word for each core document for the citing population. The composite of all citation contexts for the specialty is used as the reference corpus since we are looking for distinctive usage with respect to the specialty as a whole.

For example, the word with the highest log likelihood for the document labeled Burgi-2003 is “contact” and the second highest is “resistance”. The log likelihood uses the expected rates of these words for the Burgi-2003 contexts compared to the composite contexts for the specialty. Seven out of the ten contexts for this document contain the word "contact", and six of ten contain the word "resistance". The phrase "contact resistance" appears in one-half of the contexts.

The documents on map (Fig. 2) have been labeled by technical vocabulary. The terms are drawn from the citation contexts for each document and have the highest log likelihood with respect to the specialty as a whole. Of course, the emergence of distinctive technical terms is consistent with the notion of highly cited papers as concept symbols (Small 1978), which has also been confirmed in the work of Schneider (2006). The most striking feature is the progression of names of materials across the map showing the range of substances that were used to create organic thin-film transistors and their clustering in different sub-regions. Progressing from the lower left and moving clockwise, the organic materials include: pentacene, oligofluorenes, oligothiophenes, sexithiophene, tetrathiofulvalene, and rubrene. On the right side, some of the core papers depart from the materials theme, and deal with devices such as flexible displays, dielectrics, and inkjet printing. For about half the core documents, the term with the highest log likelihood was the same as the technical word appearing in the highest fraction of contexts for each document, a measure termed "percent uniformity" in a prior study (Small 1978).

The same method can be used to label regions of the map by grouping core documents and their contexts. For example, grouping the contexts for the core documents on the lower left gives the key term "p-type", which signifies a hole-conducting semiconductor. The key term for document contexts at the upper left is "n-type" in which electrons are conducting rather than holes.

Sentiment analysis of document co-citation links

Since we have no way of knowing what sentiments might be relevant to this specialty, we are faced with a chicken and egg situation. We need to look at the keyword data in order to hypothesize that a sentiment may be important, but this knowledge should not bias our analysis. To avoid this bias we can use a thesaurus to expand the sentiment definition by use of synonyms, and also use both the target and comparison corpora to build the word set for the sentiment. It is necessary to use sets of synonyms and related terms, rather than single words, since there are many ways a given sentiment can be expressed.

The sentiments that emerged from the initial examination of the keyword data are given in Table 2. Each sentiment is associated with a set of cue words (Finney 1979) that are searched for in the contexts and generate a count for the sentiment. Of course, relying on words to have the same meanings across varying usage contexts is risky, so this uniterm matching should be regarded as an approximation to a more sophisticated matching method which would use word pattern searching. Since we are concerned here with the numerical preponderance of cases, rather than individual instances, the errors introduced by word sense variations are less a cause for concern.

Rather than investigate sentiments for core documents, which is of course possible, the analysis focused on the co-citation links. Hence, we are looking at contexts that provide internal cohesion for the specialty, in contrast to interdisciplinary links. First, a minimal spanning tree is defined through the network of 26 nodes, including the links with the strongest cosine coefficient for each node. Sets of co-citing contexts are identified for each of the 25 links along the path. For each set of contexts, a keyword analysis is carried out, focusing not on technical words, but on the sentiment profiles defined above.

Figure 3 shows the maps with some of the links labeled by sentiment. For example, the links at the lower left are strong for the sentiments of "difficulty" and "modification". The

Table 2 Sentiments for within-specialty citation contexts

Sentiment	Sample cue words
Importance	Best, crucial, fundamental, good, ideal, novel, remarkable …
Discovered	Corroborated, demonstrated, established, elucidated, showed …
Achieved	Accomplished, able, capable, ability …
Improved	Enhance, improvement, better …
Applied	Employ, apply, use …
Modified	Alternative, approaches, attempt, tried, transformed …
Agreement	All, common, majority, most, typical, widely …
Compared	Comparable, contrast, similar, compare …
Difficulty	Although, avoid, but, challenging, circumvent, concern, drawback …
Interest	Attractive, intriguing, interesting, attention …
Future	Hope, opportunity, promise, future …

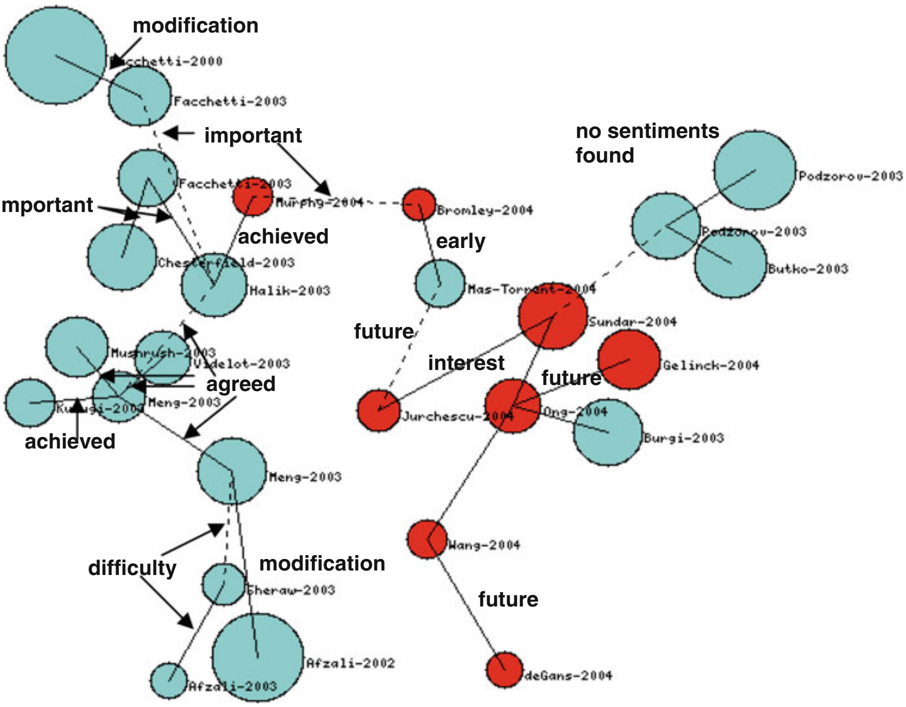

Fig. 3 Organic thin-film transistor cluster: links labeled by sentiment

technical reason for this is that certain molecular structures with high conductivity seem also to have resulted in solubility problems that needed to be overcome by chemical modification.

For example, the link between Sheraw-2003 and Afzali-2003 at the lower left labeled "difficulty" has eight co-citing contexts, four of which contain "difficulty" words. The cue words for the "difficulty" sentiment occur 5.6 times more frequently in these contexts than

Table 3 Sentiments for link between Sheraw-2003 and Afzali-2003

Sentiment	Link	Overall	Log likelihood	Significance
"difficulty"	4.5% (12)	0.83% (64)	20.0	0.000008
"modification"	3.0% (8)	0.74% (57)	9.9	0.002

Concordance examples
Difficulty
"… and thus it is *difficult* to form a smooth and uniform film …"
"… to overcome these *problems* several groups have taken the approach …"
"A common strategy to avoid this *problem* …"
Modification
"… studies have reported *modification* of pentacene to improve …"
"*Alternate* strategies that circumvent this concern …"

in all the thin-film contexts. The "modification" sentiment cue words appear 4.3 times more frequently than in all thin-film contexts. Table 3 shows word frequency as a percentage of the respective corpora, the log likelihood statistic and concordance examples for each sentiment. In addition, to show what the difficulty is about, the "difficulty" sentiment words can be linked to technical terms for "solubility" with a co-occurrence of 5 and a mean separation of 2.6 words. Thus, once a sentiment is detected it may be possible to link it to the technical terms it pertains to.

The linkages just above the "difficulty" region around Meng-2003 on Fig. 3 consist mostly of "agreement" and "consensus" terms. This appears to be due to a number of researchers having achieved comparable results in creating p-type, or electron hole conducting devices. The "importance" sentiment, prominent in the region at the upper left, is due to new opportunities offered by the discovery of organic materials capable of n-type, or electron, conductivity. Some of the links towards the middle of the map reflect temporal relations, "early" and "future". On the right, most links do not have statistically significant sentiments, except for two links on the lower right where "future" terms are indicated. The "future" sentiment appears to be associated with projected technological applications for thin-film transistors.

Redundancy, the extent to which co-cited documents share identical contexts, can be computed for each co-citation link, and it is interesting to note that the lowest redundancies fall on the right side of the map, which deals with applications and has lower sentiment values. One explanation for this is that discussion of future technological applications involve fewer competing knowledge claims and can be dealt with in a more matter-of-fact manner and with less sentiment. The left side of the map has much higher co-citation redundancy and also higher sentiment values.

Within- versus between-specialty contexts

Next we expand this view beyond a single specialty and look at the linkages that connect specialties or disciplines together and how the sentiments differ at this level. To do this we create two random samples of links: one representing links within specialties and the other, links that connect different specialties or disciplines together on a map of science.

We start with a 2009 science map (Small 2010) representing four clustering iterations nested hierarchically (Fig. 4). About 300,000 co-citing papers had within-specialty co-citations in 2,100 first level clusters contained in the higher level clusters shown on the map. The within-specialty sample consisted of 36 randomly selected co-citing papers, each representing a co-cited document pair within a first level cluster, generating a set of 67 random citing passages. While this sample is small, our aim is to generate hypotheses for testing on larger samples.

This within-specialty sample was found to cover 23 disciplines ranging from astronomy and astrophysics to psychiatry with no marked concentration in any one category. This was determined by finding the most prominent Web of Science journal categories for each first level cluster sampled.

To generate a random sample of between-specialty co-citations, we started with all the links connecting the 83 third level clusters. The map has about 1,600 distinct inter-cluster links making it 49% connected. Only the strongest links are shown (Fig. 4). About 70,000 co-citing papers are involved in the inter-cluster links. Each link on the map has on average 56 co-citing papers, and each co-citing paper participates on average in 1.3 inter-cluster links. The co-citing papers were sampled randomly and 38 were selected, generating 60 distinct contexts. The sample spanned about 46% of the 83 clusters on the map.

For the between-specialty sample, only the context for the more weakly weighted third level cluster of each pair was used. The weight for a co-citing paper is defined as the number of references cited in each cluster. The "home cluster" is defined as the one receiving the largest number of a citing paper's references. It is assumed that the "non-home" references are those that best represent the between-specialty or interdisciplinary sentiments, and thus the analysis focuses on these contexts.

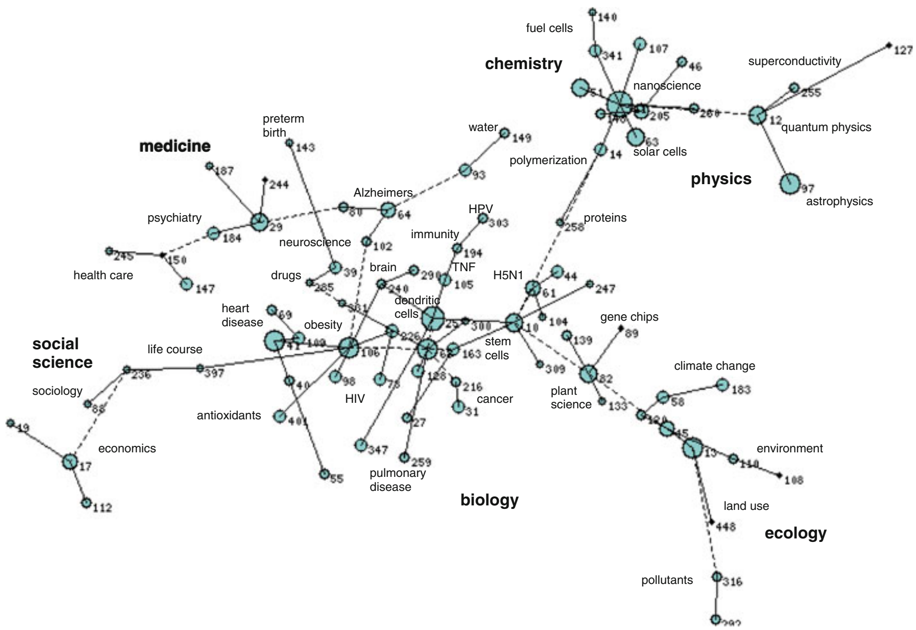

Fig. 4 Map of science from August 2009 Web of science data

Table 4 Sentiments for between-specialty citation contexts

Sentiment	Sample cue words
Importance	Significant, best, crucial, fundamental, ideal, notable, remarkable …
Utility	Employed, with, applied, used, using, utilizing, application …
Report	Described, discussed, account, stated, published, reviewed, observed …
Consensus	All, common, majority, most, typical, widely, well-known …
Uncertainty	May, might, could, should, possible, potentially, projected …
Differentiation	Contrast, differs, difference, compared …
Negation	Not, although, however, but, failed, controversial …

To show the disciplinary spread of the between-specialty contexts, Web of Science categories were compiled for the linked third level clusters, as was done for the within-specialty sample. We first determined the dominant Web of Science category for each linked cluster (at the level of core documents), reducing them to inter-category links. Of the distinct inter-category links, 80% represent pairs of different categories and 20% link the same category. A total of 18 different categories were represented on either side of the between-specialty links.

For contrasting within- and between-specialty sentiments, an initial examination of the word frequency data indicated that a smaller set of sentiments (Table 4) would suffice compared to the set used for the thin-film contexts. For the within-specialty sample, there are 67 unique contexts containing 996 words. The first technical word, "patients", appears at rank 25 with nine occurrences. The 60 between-specialty contexts contain 1,005 words. The first technical term to appear is "gene" at rank 29 with eight occurrences.

In examining the between- versus within-specialty corpora, without invoking sentiment profiles, it was found that the terms "may" and "could", so-called modal auxiliaries, had the highest log likelihood. This suggested that the notion of "uncertainty" or "contingency"—some linguists call this "hedging" (Hyland 1998)—is prominent in these contexts, and a set of synonyms and related terms was constructed. The set consisted of words such as "may", "could", "possible", "promising", "potential", etc. Charles (2006) has drawn attention to the role of "uncertainty" in the reporting clauses of scholarly texts. Here we identify a possible structural correlate.

Rerunning the analysis with this cue word set, the "uncertainty" sentiment now occurs 10 times more frequently in the between-specialty corpus than in the within-specialty (Table 5), and has the highest log likelihood ($P < 0.00002$).

Other sentiments that stand out statistically are "differentiation" and "negation", although the log likelihood scores for these were not as high ($P < 0.03$ and 0.06 respectively). "Negation" and "differentiation", like "uncertainty", can be thought of as distancing strategies.

Concordance examples for the "uncertainty" words (see Table 5) seem to support the notion that interdisciplinary citations represent knowledge that is less certain than within-specialty knowledge, thus relating a sentiment to a structural characteristic. However, the automatic profile matching could err by either assigning too many or too few contexts to the given sentiment, perhaps due to errors in word sense, or to an incomplete set of synonyms. Thus, to test the accuracy of using cue words to automatically assign sentiment, we can compare it to a manual classification. To do this the analysis must shift from individual words to full contexts since looking at each word in context is too laborious.

Table 5 Between-specialty versus within-specialty citations

Sentiment	Between	Within	Log likelihood	Significance
"uncertainty"	1.02% (24)	0.13% (3)	18.3	0.00002
"difference"	0.38% (9)	0.09% (2)	4.7	0.03
"negation"	1.45% (34)	0.87% (20)	3.5	0.06

Concordance examples
Uncertainty
"Using lithography, it *may* be possible to reduce…"
"… variants of TCF7L2 *could* be acting through…"
"… have suggested a *possible* role for CCR5…"
Differentiation
"In *contrast*, CTLA4Ig has been studied…"
"… showed significant *differences*…"
Negation
"… *but* previous studies suggest…"
"The *controversy* over whether this effect…"

The procedure is to have an indexer examine each context and decide whether or not it fits the "uncertainty" definition based on a general description of the category, but without knowing what cue words were used to make the automatic assignment. For the automatic assignment, we count the number of distinct contexts containing one or more of the sentiment words. Then a comparison is made between the human and automatic assignments.

An experiment with one indexer gave an agreement of 0.65 using Cohen's Kappa (1968) for the automatic and manual assignments of the between- and within-specialty samples. Using only the manual classification for the two samples, assuming it to be more definitive, also gives statistically significant log likelihood and chi square tests ($P < 0.0001$). Thus, a greater concentration of "uncertainty" sentiment words is found at the interdisciplinary level for both automatic and manual classifications.

Comparing the corpora the other way around (Table 6), that is, for sentiments prominent in the within-specialty set compared to between-specialty, the most striking is an expression of "use" or "utility". In other words, within a specialty the researchers are more likely to be using concepts, data, tools, and methods in an instrumental manner. Thus, within-specialty citations have a strong procedural character. This is reminiscent of Bridgman's theory of operationalism, that science consists of a series of concrete operations (Losee 1972), and also seems consistent with Kuhn's notion of normal science as puzzle solving (1970), as well as Latour's inscription devices in the laboratory (1979, Chap. 2).

Table 6 Within-specialty vs. between specialty citations

Sentiment	Within	Between	Log likelihood	Significance
"utility"	2.95% (68)	1.53% (36)	10.8	0.001
"importance"	1.0% (23)	0.51% (12)	3.7	0.05
"reporting"	1.34% (31)	0.81% (19)	3.1	0.08

Concordance examples
Utility
"… studies *with* at least nine other agents have…"
"HPV testing by *use* of clinically validated…"
"… *employing* a standard glucose tolerance…"
Importance
"… with the *remarkable* discovery of…"
"… based on a *novel* algebraic structure…"
Reporting
"It was *reported* that…"
"… Sequences were *described* previously…"

Other within-specialty sentiments that stand out at lower levels of significance are "importance" and "reporting" (Table 6). Researchers are slightly more likely to label cited work as important within the specialty, perhaps showing a tendency towards evaluative statements. The higher incidence of "reporting" terms (Swales 1990, p 150) may also be due to a greater concern with recognition within the specialty.

The between- and within-specialty samples can also be compared with the thin-film specialty contexts. In this subject specific comparison, we expect to see ways in which thin-film work is distinctive relative to these broader samples. Indeed, when we compare the thin-film contexts against the between-specialty contexts, a new sentiment appears which could be called "production" reflecting the activity of making or fabricating devices (terms such as "made" and "obtained").

Hub versus spoke analyses

Another kind of comparison involves sampling interdisciplinary contexts that emanate from a single cluster like the spokes of a wheel. We call this hub versus spoke analysis. The area selected as the hub is the third level cluster on nanoscience, which contains the thin-film transistor specialty as one of its major sub-clusters.

The nanoscience interdisciplinary sample consists of four sub-samples representing significant links to nanoscience: cancer research, immunology, climatology, and a composite of physical science areas. Hence the spread of this sample is quite broad as shown by the dotted lines on Fig. 5. The sample consists of 65 contexts having 1,081 distinct words. The most frequent technical word in this sample is "organic" with 10 occurrences at rank 26.

The link from nanoscience to cancer research deals with technologies such as microfluidics, fluorescence probes, and hydrogels which assist in diagnosis or treatment. For immunology there is interest in biological sensors and polymer-based nanofibers and scaffolds for growing stem cells. In climate research, nanoscience applications involve atmospheric aerosol reactions, and sensor technology for climate monitoring. On the other hand, the nanoscience links to other physical science areas are concerned with topics of mutual interest to the linked areas, such as organic chemistry and electronics.

In comparing the interdisciplinary nanoscience sample against the generic between-specialty context sample, only the "importance" sentiment had positive log likelihood at $P < 0.09$, suggesting that "importance" is attributed to the interdisciplinary nanoscience citations slightly more often than to interdisciplinary contexts in general.

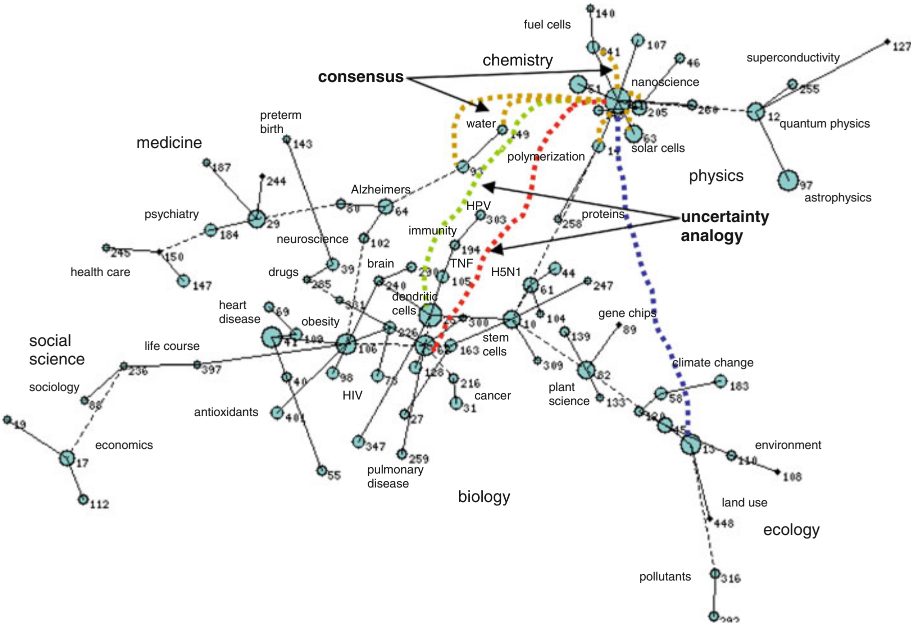

Fig. 5 Interdisciplinary links from nanoscience

Table 7 Nanoscience to physical science interdisciplinary citations versus thin-film within-specialty and between-specialty citations

Sentiment	Nano to physical sci.	Within thin-film	Log likelihood	Significance
"consensus"	2.35% (16)	1.28% (98)	4.4	0.03
Sentiment	Nano to physical sci.	Between-specialty	Log likelihood	Significance
"consensus"	2.35% (16)	1.15% (27)	4.8	0.03

Concordance examples
"…overview about nearly *all* types of FCs…"
"… Energy transfer is *widely* employed…"
"One of the *well-known* events…"
"… for a *variety* of applications such as…"

However, sentiment differences emerged most clearly by disaggregating the interdisciplinary nanoscience links. For example, for the subset of citations from nanoscience to other physical science areas, the "consensus" sentiment emerged as significant with respect to both the within-specialty thin-film sample and the between-specialty sample (Table 7). Swales (1990, p 144) calls these type of statements "centrality claims". Our finding suggests that nanoscience is engaging other physical science areas in a consensual manner (Fig. 5), focusing on what they have in common, rather than how they are different, that is, how nanoscience is consistent with other physical science areas.

Another subject specific finding comes from the sub-sample of links from nanoscience to immunology. In this between-specialty sample, as before, contexts expressing

Table 8 Concordance examples for the analogy sentiment in nanoscience to immunology interdisciplinary citations

"… structures that can *mimic* the fibrous texture of collagen …"
"… very few examples of *analogous* small molecule …"
"*Analogous* protein arrays produced via antibody adsorption …"

"uncertainty" were more prevalent than in the within-specialty sample ($P < 0.001$). In addition, there is a weakly significant tendency ($P < 0.09$) for "analogy" terms ("similar", "mimic", "analogous") to appear more frequently in nanoscience to immunology contexts relative to the within-specialty contexts (Table 8 shows concordance examples, and links are labeled on Fig. 5). Technically this is due to efforts to design nano-structures to mimic biological structures, and is consistent with the results of an earlier study on the role of "analogy" in interdisciplinary links (Small 2010).

Conclusions and future directions

The goal of this work was to see whether citation context sentiments, defined broadly as attitudes or dispositions towards cited work, can be used to characterize various regions on maps of science, both to assist in their interpretation and to discover how sentiment is related to structure. To summarize the approach, specialty or global maps are used to define sets of citation contexts to compare in terms of sentiment and technical content. These comparisons can be subject specific as in the case of the thin-film specialty, or generic as in the case of between- and within-specialty contexts. The corpora are analyzed in terms of key words using the methods of corpus linguistics.

Sentiments suggested by a preliminary examination of the data are fleshed out using word synonyms and related terms to define a cue word set, taking into account the words in both the target and comparison corpora. The sentiment words are analyzed as a group, and statistical significance is determined. Findings can also be tested by a manual classification of the contexts and compared against the automatic classification, moving from a word level analysis to individual contexts for verification.

In this exploratory study, a number of sentiments were found to be associated with structural features of maps at the specialty and interdisciplinary level, providing hypotheses for further testing. The main limitation of the present study is the small sample sizes used and the lack of a comprehensive system of sentiment categories to guide the analysis. Future work with access to full text in electronic form is clearly the next step.

In our specialty case study we found that citation context sentiments vary over the specialty map, and are prominent in regions where there is competition among researchers. In most cases, it was possible to interpret the sentiments in terms of technical or social factors. Within a specialty we were able to show sentiments for the resolution of difficulties, modification of methods, identification of important discoveries, areas of agreement, and prospects for future developments. Thus, sentiments provide insights into the current issues and concerns of a research community.

Comparisons of within- and between-specialty citations revealed a number of sentiment differences. For between-specialty contexts, sentiments of "uncertainty", "difference" and "negation" appear to be important, while for within-specialty contexts, sentiments "utility", "importance" and "reporting" are prominent. In interdisciplinary science, "uncertainty" can be interpreted as scientists reaching outside their own fields for ideas and solutions.

Within a specialty, "utility" may express a focus on procedure and process, and "importance" may be an expression of the socially motivated need for evaluation.

Finally, sentiments for interdisciplinary contexts emanating from a hub were seen to depend on proximity to the hub and the nature of the linked topic, for example, the "consensus" sentiment for nearby topics, and "uncertainty" or "analogy" for topics further removed. The "consensus" sentiment was interpreted as reflecting solidarity with neighboring areas, while analogy is seen as strategy for finding novel solutions by drawing on similar problems in other fields. Thus, citation structure and sentiment appear to be related in a number of ways, opening up new ways to explain structure.

In future research it should be possible to connect sentiments to the technical terms they apply to by a word pairing and proximity procedure, linking sentiment terms to technical terms. This would tell us, for example, what the "difficulty" or "uncertainty" is about. In addition, we need to study how sentiments evolve over time, how they differ from one specialty to another, and from one interdisciplinary link to another. Of course, to do this, larger sets of contexts are needed to increase sample size, and to develop baseline corpora for all of science as well as individual disciplines and specialties.

References

Argamon, S., Dodick, J., & Chase, P. (2008). Language use reflects scientific methodology: A corpus-based study of peer-reviewed journal articles. *Scientometrics, 75*(2), 203–238.

Börner, K., Chen, C. M., & Boyack, K. W. (2003). Visualizing knowledge domains. *Annual Review of Information Science and Technology, 37*, 179–255.

Charles, M. (2006). Phraseological patterns in reporting clauses used in citation: A corpus based study of theses in two disciplines. *English for Specific Purposes, 25*(3), 310–331.

Chen, C., Ibekwe-SanJuan, F., & Hou, J. (2010). The structure and dynamics of cocitation clusters: A multiple-perspective cocitation analysis. *Journal of the American Society for Information Science, 61*(7), 1386–1409.

Chubin, D. E., & Moitra, S. D. (1975). Content analysis of references: Adjunct or alternative to citation counting? *Social Studies of Science, 5*, 423–441.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scale disagreement or partial credit. *Psychological Bulletin, 70*, 213–220.

Cozzens, S. E. (1985). Comparing the sciences: Citation context analysis of papers from neurophramacology and the sociology of science. *Social Studies of Science, 15*(1), 127–153.

Finney, B. (1979). The reference characteristics of scientific texts. Unpublished Master's Thesis. The City University of London.

Hargens, L. L. (2000). Using the literature: Reference networks, reference contexts, and the social structure of scholarship. *American Sociological Review, 65*(6), 846–865.

Harwood, N. (2009). An interview based study of the functions of citations in academic writing across two disciplines. *Journal of Pragmatics, 41*(3), 497–518.

Hyland, K. (1998). *Hedging in scientific research articles*. Amsterdam: John Benjamins Publishing Company.

Khan, K., Baharudin, B.B., Khan, A. & Faxal-e-Malik. (2009). Mining opinion from text documents: a survey. In *2009 3rd IEEE International Conference on Digital Ecosystems and Technologies* (pp. 194–199). New York: IEEE.

Klavans, R., & Boyack, K. W. (2009). Toward a consensus map of science. *Journal of the American Society for Information Science, 60*(3), 455–476.

Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.

Kuhn, T. S. (2000). The road since structure, Chap. 4. Chicago: University of Chicago Press.

Latour, B., & Woolgar, S. (1979). *Laboratory life: The social construction of scientific facts*. Beverly Hills, California: Sage.

Losee, J. (1972). A historical introduction to the philosophy of science. London: Oxford University Press.

Moravcsik, M., & Murugesan, P. (1975). Some results on the function and quality of citations. *Social Studies of Science, 5*, 86–92.

Paquot, M., & Bestgen, Y. (2009). Distinctive words in academic writing: A comparison of three statistical tests for keyword extraction. *Corpora: Pragmatics and Discourse, 68*, 247–269.

Schneider, J. W. (2006). Concept symbols revisited: Naming clusters by parsing and filtering of noun phrases from citation contexts of concept symbols. *Scientometrics, 68*(3), 573–593.

Scott, M. (2010). *WordSmith Tools Step by Step*. http://www.lexically.org/wordsmith/step_by_step_guide_English.pdf.

Small, H. (1978). Cited documents as concept symbols. *Social Studies of Science, 8*, 237–340.

Small, H. (1986). Synthesis of specialty narratives from co-citation clusters. *Journal of the American Society for Information Science, 37*(3), 97–110.

Small, H. (2010). Maps of science as interdisciplinary discourse: Co-citation contexts and the role of analogy. *Scientometrics, 83*, 835–849.

Small, H., & Upham, P. (2009). Citation structure of an emerging research area on the verge of application. *Scientometrics, 79*(2), 365–375.

Swales, J. M. (1990). *Genera analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Teufel, S. (2010). *The structure of scientific articles: Applications to citation indexing and summarization*. Stanford, California: CSLI Publications.

Verlic, M., Stiglic, G., Kocbek, S. & Kokol, P. (2008). Sentiment in science: a case study of CBMS contributions in years 2003 to 2007. In: *21st IEEE International Symposium on Computer Based Medical Systems* (pp. 138–143). New York: IEEE.