An area concentration exam in bibliometrics and computational linguistics

Introduction

The following essay sets out what I see as the primary technical competencies in compiling a conceptual history of the “classics” in sociological theory. This conceptual history will consist of a catalog of all published interpretations of core theoretical concepts in more than twenty thousand published works in four sociological journals, using a combination of quantitative and qualitative interpretive methods. Unlike previous work in the history of sociology, this conceptual history will be machine-readable. The final dataset, the conceptual history, would supplement characterized interpretations with when they occur; the social position of the author; a formal citation structure which both links interpretations together and indicates what authors had read, or at least read about. More ambitiously the dataset would indicate the logical place of the interpretation in the argument of the author and a multitude of logical relational data connecting interpretations. Similar to Randall Collin’s *Sociology of Philosophies* (1998), this history will act as source material for theories of theoretical concept selection, development, mutation, and use in the institutionalized academic field of Sociology. Because the conceptual history will be machine-readable, I will be able to develop and test meta-theories explicitly and reproducibly.

In studying these essentializations I have chosen published academic writing as my object. Of course, theory is not only used, presented, or digested through academic publications. It is taught in academic institutions, discussed in small groups of excited intellectuals, or presented at conferences. Indeed, traces of the influence of theory often remain only implicit in a researcher’s work, possibly never reaching the surface text. Thus an exegetical dissection, however vast, must be limited in its ability to trace a wholistic development of academics’ ideas of society, and of their influence. Despite this, written text influences and is influenced by theory, and offers a highly structured precipitate for the researcher of the development of science (myself). Availability alone makes academic discourse a great candidate for analysis. JSTOR has enabled this endeavor by providing me with high-quality computer-readable transcriptions of nearly every full research article, review article, discussion, correction, book review, editorial, and introduction from American Journal of Sociology, American Review of Sociology, American Sociological Review, and Social Forces from the late 1800s to present. Many papers in the dataset cite the classics simply for its use in their arguments or uses concepts without citation. Others consciously dissect, reconstruct, and reinterpret the classics. This vast body of academic literature will be the primary source from which I construct my conceptual history. I am able to observe the full social act of the academic statement in exactly the form it occurred, something which much of sociological analysis cannot say of its object of study.

In the following I begin with a short motivational section which surveys a modicum of the questions and proposed answers in the sociology of knowledge and sociology of scientific knowledge, as a justification for and framing of this study. I then survey the two broad methodological approaches which figure centrally into this endeavor. First, bibliometric analysis offers techniques for tracking the flow of ideas through a network of citations, for identifying the composition of intellectual sub-groups over time, among other things. Second, literature in computational linguistics and linguistics proper offers the computational and theoretical tools to classify and relate the discursive acts comprising these academic articles on such a large scale.

Motivation and theoretical inspiration

This project is a precursor to understanding the development of the classics into and as classics, and how theoretical concepts and modes of reasoning have been extracted, essentialized, and reformulated from these works. In this section I present some meaningful questions which can be asked empirically of sociological literature and its relation to the classics. I do not attempt at an evaluation of the credibility of propositions, theories, or lines of research, and intend this section as an embodiment of the consensus of the value in the classics themselves. That is, that they at least give us good ideas (Fine and Kleinman 1986). Understanding the mechanisms underlying the development and dissemination of classical theoretical concepts in Sociology has enjoyed quite a bit of exegetical analysis to date. Baehr's *Founders, Classics, and Canons* (2016) gives the most thorough account available of past theories along these lines, reviewing over fifty years of inquiry into the subject. Some central questions of the book are why these specific works have become classics, why we still read the classics in sociology and if we should, and how the classics are used in contemporary works. Much of my discussion in this section, if not directly influenced by Baehr's account, were at least echoed again by it after my independent discovery.

In all scientific research there is a living dialectic between theory and reality, as what is observed impinges on possible theories, and theories then direct the academic mind towards new research (e.g. Parsons 2010:9). Yet the uniquely central place of the founders of sociology, as "canonized," makes the field an exceptionally interesting case. The role of theory as placing and legitimizing research in sociology, and as justifying particular analytic moves, gives these essentializations a particularly powerful directive force. Such a strong institutionalized requirement for reasonable theoretical framing suggests functional explanations to what ends up being referenced in this justification of the paper, not to mention the interpretations which are commonly employed, possibly with some independence from the original argument and its purposes.

One fruitful angle is to consider what is constructed of the classics as a shared language, what Peter Baehr argues is one of the redeeming functions of the classics in sociology (Baehr 2016:81). The classical arguments are taught to all sociologists in their institutional training, and subsequently reinforced through their disproportionate attention in the front ends of papers. These shared concepts allow complex stances, dichotomies and causal frameworks to be referenced without redressing all the details (Baehr 2016:82). A similar functional role for the classics is argued by Nicos Mouzelis (1995:8), that sociological theory can act as a lingua franca, allowing subdisciplines of the social sciences to communicate with one another. In that the classics constitute a shared language of the discipline one could theorize them as a framework without which sociological thought could not occur, and which color all its manifestations. In light of this, a dissection and philology of the conceptions and essentializations of the classics carries a potential for uncovering the dynamics of the very framework of sociological thought.

This view is confronted by various charges of the classics being ambiguous and their interpretation being contentious. For example, the logical inconsistency of Mead's *Mind Self and Society* spawned two conflicting schools of interpretation, the Revisionists and Blumerians, who indeed didn't read Mead incorrectly, only selectively (Fine and Kleinman 1986:135). Even more drastic, Masterman (1970:61–65) identifies a full twenty-one different senses of ‘paradigm’ in Kuhn’s *Structure of Scientific Revolutions*. Davis (1986:295–96) argued that a crucial requirement of a lasting classic in sociology is precisely this ambiguity, the ability of a work to be interpreted in a variety of ways by subsequent researchers for many different reasons. There is even some debate in the question of from what perspective we should interpret the classics. On one side the historicists argue that for a “true” interpretation you must understand the author’s social and intellectual situation, including the positions and writings of contemporaries with which the author is addressing. On the other side, presentists argue that this endeavor should be abandoned as useless or impossible, and literature should be instead read for the purposes of now, as if the author were speaking directly to the modern reader. Either stance leads to its own interpretive nightmares, and the existence of proponents on both sides further suggests heterogeneity in interpretations of any classic text. For an excellent and thorough account of this debate, see Baehr (2016:93-105).

This dichotomy – between the classics acting as some lingua franca and the interminably vague nature of understanding that which has been written in another time to a foreign people – suggests many interesting theoretical questions which could be asked to the data. How does this ambiguity show itself in uses of the classics in sociological research? Is there a small discrete set of interpretations, or a vast continuous manifold of interpretations and uses, changing over time? How small, or how vast? Can we see the same terms being used, but with distinct places in argumentative discourse, setting off an interpretive telephone game where the research community constructs and reconstructs their theoretical underpinnings relatively freely? Do these references actually act as a shared language in some definable sense, which helps communication across sub-disciplines of sociology or of social science more generally, or do varied interpretations yield only confusion, rendering these theoretical touchstones as not much more than window dressing on an empirical analysis? If they do act as a common language, how does a discipline go about constructing a lingua franca from a literature which doesn’t have a unique interpretation? Is the heterogeneity in interpretation itself unevenly dispersed across theoretical concepts, and if so what does this map look like?

Quantitative analysis versus qualitative analysis

With such potential theoretical and empirical complexity afoot, this study will begin deductively, attempting at an honest depiction of what has been taken, used, and understood from classic sociological texts. I do not intend to formulate theories before looking carefully at the source material. I have the great fortune in this methodological choice to be led by the source material itself. Parsons tells us that "true scientific theory is not the product of idle 'speculation,' of spinning out the logical implications of assumptions, but of observation, reasoning, and verification, starting with the facts and continually returning to the facts." (Parsons [1937]2010:vi). Similarly, Blumer implores the researcher to have respect for the "obdurate character of the empirical world" (Blumer 1969:23). Blumer felt that social scientists held respect only for the "sanctioned scheme of scientific inquiry," such that they felt no need for firsthand contact with the sphere of life they were describing, their results thus reflecting the method itself rather than the object of study (38). Geertz implores the researcher against seeking generalizations, for example claiming that any general definition of religion must necessarily be vacuous (1973:40) and sees the general as much less typical of humankind than the particular. He implores that a deep and true understanding of the specific is what is needed to come to know the underlying processes more generally (43-44). I recognize it may be problematic that I include such classic theorists as buttresses in my analyses but am somewhat comforted by the fact that I have read these books in their entirety, and that their empirical work has had lasting significance in the social sciences. Nonetheless I feel I will have to return to this section after diving deeply into others' practices of interpretative justification in some sort reflexive analysis.

In light of an emphasis on a focus on the phenomenon, on the complexities of the individual instance, a relevant question to ask before diving too deep into bibliographic and computational linguistic methods is to what extent they will mislead, sending me down methodological rabbit holes instead of showing me what is empirically present in the content itself. I cannot avoid the fear that every hour I spend working with black-box linguistic algorithms is an hour I could have been absorbing the literature itself, possibly only deferring interpretation to a later stage, at which time it is likely to be shallower than it would have been without the computer screen getting in the way. As in many areas of social research, computational methods trade scale for nuance, and an ethnographic analysis should be preferred to a quantitative one, when it is feasible.

Playing with this fundamental question, I took cursory look at my dataset of published sociological work. I've counted simple word frequencies of some major concepts arising from the source literature in AJS, ASR, ARS, AJPS, ASQ, and Social Forces, comprising 22,283 articles, summarized in Table 1. These few concepts already constitute a formidable task for qualitative coding, although such large-scale qualitative analyses do occur in the literature (e.g. Lin 2018). Once concepts become crystallized, very many of these references may prove redundant, vastly reducing the number I would need to analyze qualitatively. Quantitative methods will prove instrumental in this reduction process, however, giving me an indication of the groups of similar interpretations, and pointing me to outliers I would not have found if coding some random subset of the data. In addition, methods for focusing my analysis, such as main path analysis, may reduce the number of articles under study such that I can read them in full. Thus the short answer which will guide me in the following review is that quantitative methods help me to focus a qualitative analysis, and in some cases can extend interpretative methods I develop on a small nonrepresentative sample to the entire population.

Term	Year of first mention	Number of articles with at least one mention	Total number of mentions	Average number of mentions per mentioning article
taken for granted	1895	918	1197	1.30
the stranger	1896	188	603	3.21
dialectical	1899	531	1419	2.67
organic solidarity	1899	143	359	2.51
social capital	1910	1154	9922	8.60
significant other	1930	128	1037	8.10
protestant ethic	1931	392	951	2.43
structuration	1934	303	726	2.40
symbolic interactionism	1951	336	868	2.58
discrepant roles	1954	3	4	1.33
dramaturgical	1962	92	178	1.93
normal science	1966	128	229	1.79
thick description	1976	137	168	1.23
structural hole	1993	31	691	22.29
matrix of domination	1995	7	10	1.43

Table 1: A simple count of mentions of several theoretical concepts originating in classic sociological works. The counts are only among the 22,283 documents in my dataset tagged by JSTOR as type *research-article*. This excludes 41,421 documents of type *review-article, in-brief, discussion, correction, book-review, editorial, introduction, index, news, misc,* and *other*. The table includes a total of 17,165 mentions to the 15 concepts. Notice the preponderance of mentions of structural hole within the mere 31 publications which mention it. These counts are likely underestimates, as they only include instances where the word or phrase is followed by a non-alphanumeric symbol, such as a space, apostrophe, or period.

Bibliometrics

Bibliometrics, also known as citation network analysis, takes as its object of study the directed graph obtained from the bibliographies of academic works, patents, court cases, or other contexts where documents meaningfully reference each other. Library science and information retrieval scholars use such networks to map academic fields either to produce efficient and useful search results (e.g. Volanakis and Krawczyk 2018) or to produce comprehensive literature reviews (e.g. Belter 2016; Liñán and Fayolle 2015). Social scientists have studied the same data structure, developing and testing macro theories of the organization and development of academic fields (e.g. Shwed and Bearman 2010). More extensive even is the impact of physicists and computer scientists in this area. In developing methodologies for the analysis of communication networks, circuits, power networks, and other systems, they also gained an uncanny ability to describe the topology, dynamics, and history of publications in their own fields (my favorites are Kuhn, Perc, and Helbing 2014; Perc 2013; Xie et al. 2015). In addition to the authorship of citation network analyses being quite interdisciplinary, the readership of such methods is also diverse. This is because of the utility in characterizing and understanding any scientific field using bibliometric analysis. As an illustration, there are 207 papers listed in Web of Science with the topic "citation network analysis" published in 2019, coming from 149 distinct journals. In this section I will give a limited account of bibliometric research, focusing on papers directly relevant to my own project.

Formally, a citation network is a directed graph with publications as nodes, and citations in the bibliography or the body of the publication as edges between publications. In this essay I will consider an edge as going from the paper which is cited to the paper which makes the citation. Although I find this unnatural, it is the norm in the field. The nodes in a citation network will be annotated with the meta-information of *author, journal, publisher, year,* and *institution*, along with other information one can extract from the documents themselves. Note that some metadata may only be available if we have the papers themselves, such as *institution*, and may not be retrievable for nodes which are cited. There are many useful derivative networks the researcher can construct from a citation network using this metadata. We can collapse the network on any one of these attributes, either in the citer or cited articles, yielding a new network potentially of use to the researcher. For instance, we might consider institutions citing authors, journals citing other journals, or years citing years. Collapsing on any of these attributes changes the focus of analysis and the questions the researcher can ask to the network.

Another useful and widespread method of generating derivative networks is to create 1-mode projections (Breiger 1974), considering citer and cited as two separate types of nodes. These one-mode projections are commonly referred to as the co-citation network and bibliographic coupling network. The co-citation network draws an undirected weighted edge between nodes to the extent that they are cited together in the same paper. The co-citation network focuses on the intellectual basis of the papers in the dataset, and clusters in this network indicate co-use in later work. This method is quite useful for extending the citation analysis beyond the papers for which bibliographies have been gathered. Similarly, a bibliographic coupling network links citing papers to the extent that they cite the same papers. This network focuses on the papers for which we have bibliographies, and proximity or embeddedness indicate communities who read and reference the same works. Papers which have an exceptionally strong link in the bibliographic coupling network work with the same intellectual context and are likely to be on very similar topics. Both co-citation and bibliometric coupling clusters can indicate topical similarity and can aid in mapping the conceptual landscape of a field of study (Chen 2006; Chen, Ibekwe-SanJuan, and Hou 2010).

We must be careful in drawing these projections, as papers can only cite papers published before its own publication date, and thus we should only compare citation links which could have occurred, where the date of the citer publication is after the date of the cited publication (this was pointed out to me in Zhou et al. 2019). Also note that citation networks are necessarily incomplete, much in the same way social networks are. There will always be works for which we do not have the bibliography, and we cannot know of citations which occur in documents for which we do not have bibliographies. The researcher must be careful, keeping in mind these bounds in the planning and execution of an analysis.

The typical bibliometric analysis of a field, published in a journal other than *Scientometrics*, employs a very simple analysis. First it will identify the most productive authors and institutions, as well as the most cited papers and journals. It then uses a visualization software such as VOSviewer to examine the co-citation or bibliographic coupling network, usually using some cluster analysis to paint the broad strokes of the various distinct research communities in the sample. In the following I analyze a few slightly more sophisticated methods I believe figure directly into my own analysis. I have not attempted to give a comprehensive overview of the state-of-the-art methods appearing in *Scientometrics*, the leading journal in citation analysis proper, as it would lead this essay far afield of what is useful.

Main path analysis

A somewhat peripheral trend in bibliometric analysis particularly important for this project is main path analysis. Hummon and Doreian (1989) proposed the first and simplest incarnation of this method in a re-analysis of a history of DNA research conducted by Garfield, Sher, and Torpie (1964). Given a citation network, we first list each path from source to sink. A source in this network is a node for which we do not have the bibliography, and a sink is a node which has not been cited by other papers in our network. We then count the number of such paths which pass through each edge, thus obtaining weights for these edges. This method of counting is called search path count (SPC), and is one among multiple similar counting methods, exhaustively discussed in Liu, Lu, and Ho (2019). Once the network is weighted in this manner, to construct a main path we start from any node and choose the edge which leaves it (a citer) which has the highest weight. Each such choice, due to the construction of the weights in this network, intuitively corresponds to the citer which connects the most academic literature through this citation. In the case of the 65 publications analyzed qualitatively by Garfield et al. (1964) this method reproduced the most historically significant markers of development. Main path analysis has been used since this introduction in an attempt to identify the “path of progress” in a field. Note that there is a main path for each and every starting node, and that when these paths overlap, they will never diverge because of the deterministic way paths are chosen. Because of this we can identify multiple main paths by grouping those paths that result from various starting nodes, especially those which are considered important to the study at hand. Various modifications have been made to this method such as changing the weighting algorithm or choice function for choosing the next link in the path in an attempt to match most closely to the intuition of finding the most “connecting” citations. One common modification is to look for “global main paths” by looking for the path with greatest sum SPC. One can also modify the method to require inclusion of specific links (“key-route search”). Liu et al. (2019) gives the most comprehensive reference for the latest developments in this method.

Although there is a stellar dissection of social network analysis via main path analysis (Hummon and Carley 1993), there is no main path analysis to my knowledge which takes sociological literature as its object. The upshot of this method for my purposes is that it identifies publications instrumental to the development and integration of a specific publication into common knowledge and can be used to focus quantitative or qualitative linguistic analysis on a small subset of important papers. But as with most other methods in the social sciences we must proceed with caution. What exactly the main path is picking up is context dependent. For example, Hummon and Carley (1993) find that in a high-consensus normal science which builds incrementally on past achievements (in the case of their analysis, social network analysis), publications on the main path tend to be methodological or experimental articles. This is presumably because these works can be immediately integrated into a vast number of works, and not as much because they are the most influential in the development of the field on a conceptual level. Indeed, in the case of social network analysis, not a single publication which members of the field cited as influential showed up on any of the six main paths identified by the authors[1] (102). The method can be illuminating, but we should be very clear about what our method is picking up, and careful with the interpretation we paste on its output.

Measuring coherency and detecting clusters

I will now introduce recent operationalizations of the cohesion, or so-called “coherency”, of a directed network. This discussion will then dovetail nicely into a presentation of some clustering methods for directed graphs. How do we quantify the extent to which a citation network or a portion of one is integrated as a community of discourse, or conversely is disconnected and fragmented across separate communities? I was introduced to a collection of measures answering

[1] The names the authors gave for these main paths were role analysis, methods, network data, biased networks, structure, and personal networks.

this question by a recent study in AJS which found a positive correlation between ambiguity in the abstract of a research paper and the cohesion of the citation network emanating from it (McMahan and Evans 2018). The paper introduces two measures of cohesion. The first, which the paper does not use because it is "computationally expensive" (880) was developed under the term "graph entropy" in the physics community by Corominas-Murtra et al. (2010). They define the entropy of a directed network as the uncertainty one would have in tracing paths through the citation network (from more recent papers backwards in time). That is, if a path is chosen through the citation network (to a citer, then a citer, etc.), what would be the uncertainty of the exact path given only the final paper in this path? The concept is best illustrated in the three panels of Figure 2, drawn from McMahan and Evans (2018). In the far-left figure the entropy is zero because given a final node (drawn at the top of the figure), we can be certain which path was taken to get there. As we move to the right the entropy increases as there is a more variety in the possible paths taken to this final node. This method is intuitively pleasing, as if a citation network is split into separated communities there would be less uncertainty in tracing these paths backwards, whereas if the citation network is entirely integrated the uncertainty would increase. To my knowledge this method has not been used in the social sciences, and I am not entirely sure why McMahan and Evans present it at all, except that they can use the word "entropy" when presenting their results.

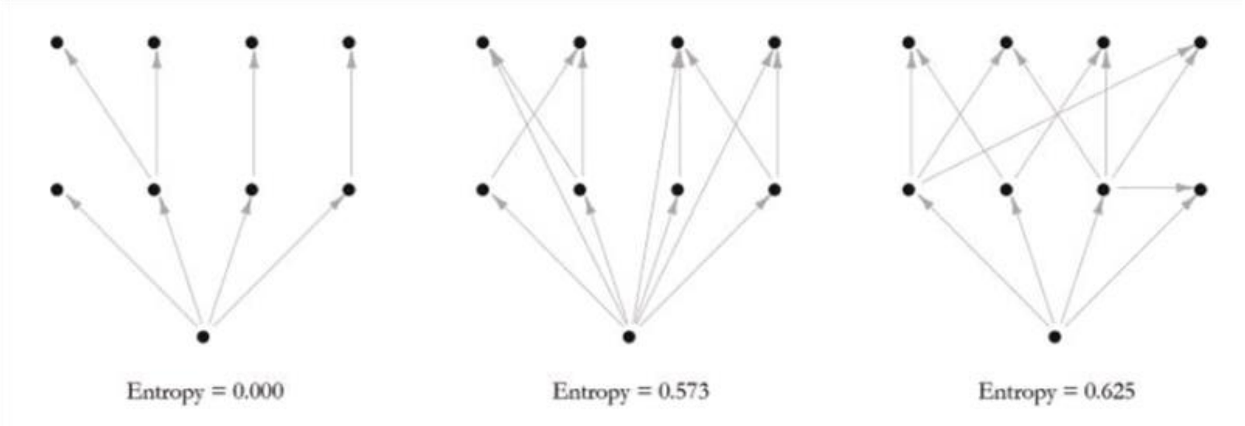

Figure 2. I have reproduced here an illustrative diagram Shwed and Bearman (2010) used to explain the topological fragmentation measure of Corominas-Murtra et al. (2010).

The second operationalization, the one which was actually used in the analysis of McMahan and Evans, is typically attributed to Newman and Girvan (2004) and is called *maximal network modularity*. The method efficiently detects communities in an undirected network and works by determining a partition of the network which maximizes “modularity”. Modularity here is the number of ties which occur within the same group minus the number of expected such ties if ties were distributed randomly. For the purpose of characterizing the extent of coherency it is not the partition itself which concerns us, but the maximal network modularity achievable by a partition, indicating to what extent the network can be partitioned into separate clusters. Although this method is more efficient than the graph entropy method cited above, it disregards the directed nature of the citation graph, and thus will in some cases generate unreasonable estimates of coherency. McMahan and Evans (2018) follow citations forward in time for each paper under consideration up to three levels deep and 1,000 papers total, generating an ego-network for each paper, and find that for these networks the two measures have a correlation of -0.7559. That is, with increasing entropy (increasing uncertainty in tracing paths backwards) we see decreasing maximal modularity (less of a coherent group structure). This method has seen at least one other application in the sociology of knowledge, in Shwed and Bearman's (2010) study of academic consensus over time in a few, perhaps carefully selected, topical case studies. In this treatment they take maximal network modularity as a measure of contestation and reproduce the agreed-upon estimates of when consensus formed in these debates (or whether they ever did). They note wisely that any such analysis is highly dependent on the citation practices of the field (833). Although the authors praise this method as stepping beyond the previously qualitative methods of SSK and ANT, their discussion reveals that qualitative assessments are indeed necessary, and I see this as yet another method for generating provisional typologies and characterizations which can then be subjected to more qualitative (or indeed computational linguistic) analysis.

There is such diversity in methods for clustering citation networks that I will only give a brief overview here. Malliaros and Vazirgiannis (2013) give an excellent and exhaustive survey of clustering methods for directed networks. The variety is not due to some insurmountable debate as to the "best" method, although this is part of the problem. It is instead due to a vast heterogeneity of purposes and contexts, and, as we saw above, of tolerances for computational inefficiency. For example, you may want to cluster structurally similar papers, who cite and are cited by the same papers (or papers within the same cluster). In this case blockmodeling is a natural method for clustering, as it was originally developed for identifying structural isomorphism in directed graphs, and outputs a higher-level citation structure between clusters, identifying citation habits of communities (see Doreian, Batagelj, and Ferligoj 2004 for an excellent textbook on the topic). Another method for clustering based on structural equivalence is to collapse the citation network into the co-citation or bibliographic coupling network and use traditional clustering algorithms on this truly undirected graph. One might alternatively wish to cluster together papers which cite each other, of which the Newman-Girvan method is an example. This clustering task has been accomplished with varying levels of efficiency through spectral, random walk, and information theoretic methods, among many others. Rosvall and Bergstrom (2008) demonstrate the creative energy permeating this area by clustering into groups based on the expected number of words it would take to describe random walks through the directed citation network. I have reproduced their beautiful map of the scientific disciplines' citation patterns in Figure 4. Each node here represents thousands of papers. As a final note, although the risk is great, researchers have also simply treated the citation graph as undirected, which makes available any clustering method usable for undirected graphs.

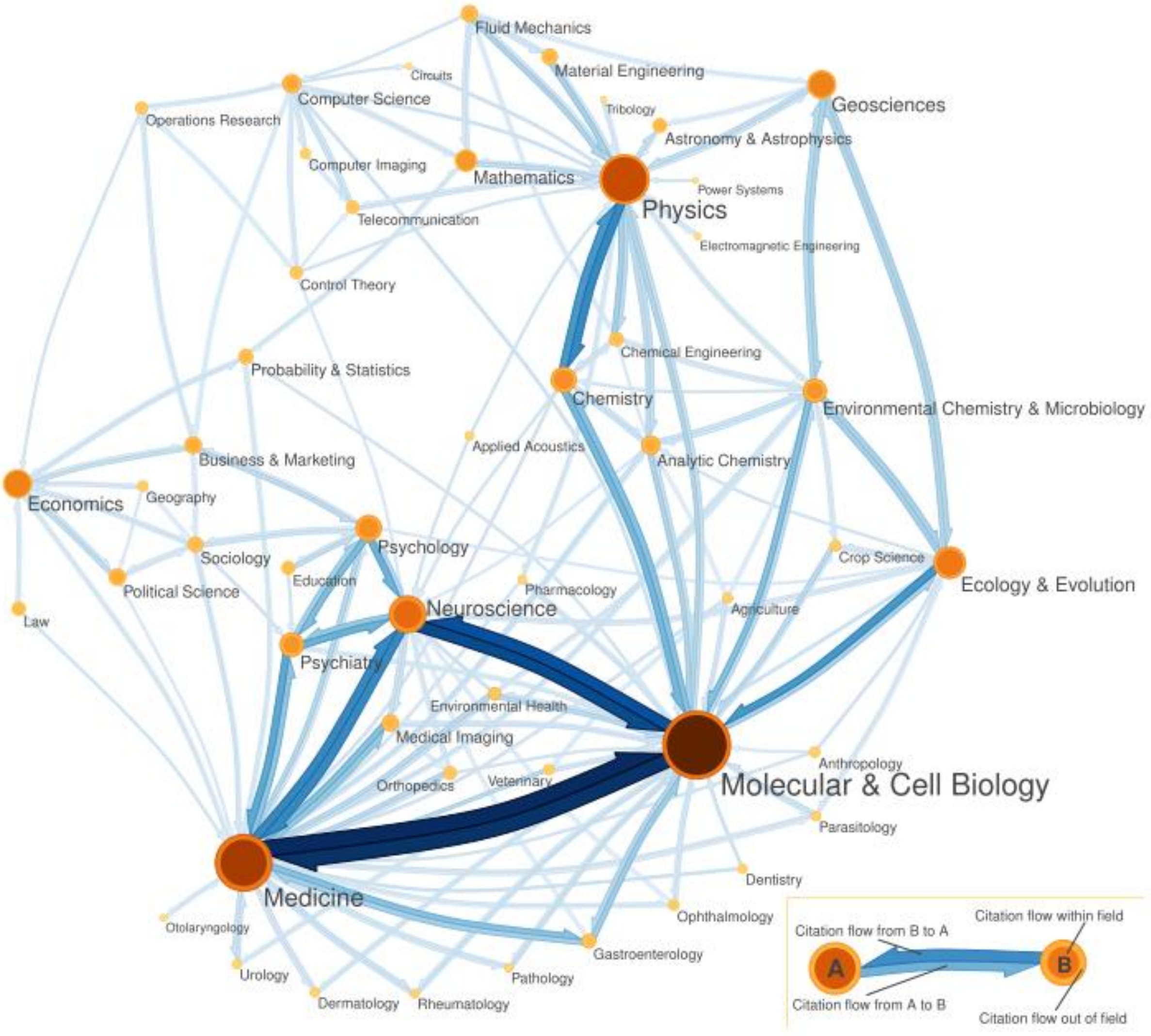

Figure 4. An absolutely stunning map of the macro-structure of citation based on approximately 6.5 million citations. This figure is from Rosvall and Bergstrom (2008)

Conclusion

Bibliometric analysis is useful for identifying pivotal publications, for pointing to the paths along which knowledge is transferred, in identifying broad shifts in academic disciplines, and in mapping the subfields which constitute any such field. I can quickly identify papers which belong to similar discourse communities and analyze how the relation between journals and institutions changes over time or identify holes where citations should probably be. It gives me a quick answer to the question "Who are these people that are citing Simmel's Conflict?", a question which transposes into “What are these people that cite Simmel’s Conflict writing about?”, a question well suited for co-citation analysis. We can examine the network as a whole to look for the higher-level intellectual structure of the discipline, or like McMahan and Evans (2018) can focus on the n-th degree “out-neighborhood” of papers who cite papers who cite … our focal paper, showing us the unwitting community surrounding each paper. In short, authors’ explicit placement of their own papers through citations gives an extremely valuable view to the researcher who is studying thousands of papers at once, directing the focus of more qualitative interpretive work, and auto-generating data-driven bibliographies of the field.

(Computational) Linguistics

Bibliometric analysis is by no means sufficient to answer the original question I posed of this literature. It is extremely useful for limiting and focusing what I read, allowing me to get right to the heart of the literature, to the most influential, the most central, to group publications for a higher-level analysis. But what does this literature say about the classics? And how, discursively, is theory incorporated into an academic culture? For this I will need to look to the content of the publications, reading and forming my own conclusions about these interpretations. Even as bibliometric analysis can focus an in-depth qualitative analysis on “important” papers, we must recognize that the structural position of papers is not the only thing interesting in them, and that there is just as much to read into what has not been cited as what has. Moreover in the flow of a publication what exactly is cited is only half of the citation, and neglects what was said about the citation and why it was cited in the first place. Furthermore, for many analyses I want to cast my interpretive net wider, seeking typologies incorporating every interpretation or understanding of sociological concepts in these texts, looking for a bird’s eye view of the development of these understandings over time. Thus it will be of great benefit to extend some parts of my interpretive process with computational linguistic methods, allowing me to automatically code interpretations of much more literature than I could read in my lifetime.

Answering critiques to bibliometric analyses

The bibliography of a paper gives a record of the intentional linking and situating an author accomplishes through their citing, but citations are embedded in their textual context, in some presentation to some audience, in a specific discussion or argument. One simple and crucial limitation of a citation analysis which considers only the bibliography is that it assumes implicitly that citations are equal in their importance and are all of the same meaning, or at least that their heterogeneity has no bearing on the conclusions of the analysis. This assumption is patently false, and this pretense unintentionally obfuscates what is really going on in the academic field under study. Substantial critiques along these lines first surfaced in the 1970s in the sociology of science, finding flaws in the use of bibliometric analyses in assessing a paper's impact (Moravcsik and Murugesan 1975) although the core empirical properties of citations which support this critique have been discussed long before, at least as early as Merton (1957). Papers with memorable titles can be highly cited without being central to the claims of the author, and indeed with only rarely being read (Moravcsik and Murugesan 1975).

The most recent and extensive effort to examine to what extent the equal-weight assumption is correct was that of Lin (2018). Lin analyzed the citation context of 25,617 in-text citations from 360 research articles across six social sciences and humanities fields. Lin found that 24.6% of citations in Sociology are perfunctory, not performing an essential role to the development of the author's thesis and mentioned only in passing. Furthermore a full 14% of citations are not contained in the main body of the text but occur in footnotes or endnotes. In an analysis exclusively of references to Weber in AJS, ASR, Social Problems, and Social Forces, Adatto and Cole (1981) found only 36% of papers used his work in a way central to the argument of the paper. These figures are hardly surprising to a reader of sociological literature, but the implication is striking, and ignored by a large portion of bibliometric researchers. That is, citations are not always informational references for the reader to reference information to better understand the paper itself. They act as signals of status and community membership or bolster claims which were not derived initially from the citations which were acquired post hoc for justification.

The answer to these critiques has been enthusiastic, as linguists and sociologists of science alike have developed theoretical typologies for classifying citations, hand in hand with the argumentative moves of an academic text more generally, in terms of the speech act their accomplishing (e.g. Swales 1986; Teufel 1999; White 2004; Siddharthan and Tidhar 2006; Jurgens et al. 2018). The speech act was originally introduced by Austin (1975) and is based on the idea that some act of communication both says something, but also *does* something. What accomplishes the act could be a statement, an entire speech, a single grunt, a paper, a section, or a sentence. Typical examples of speech acts include asking a question, challenging authority, or clarifying. Generally speaking, typologies of speech acts have proven useful for linguists in abstracting from the vast creative diversity of human communication into the intent and effect of the speech. In the case of in-text citations these speech acts are commonly called the "citation intent," and classify what the citation is doing there in the text.

There are more such typologies than need be surveyed here, so I will describe the prototypical typology given by Jurgens et al. (2018, Table 1), based previous work (Ding et al. 2014; White 2004). Simply put, in reference to a paper P, an in-text citation may say one of the following: P gives background information (background); P motivates this study (motivation); we use methods, data, etc. from P (uses); we extend P (extension); there is some similarity/difference X between this paper and P (compare/contrast); P could act as a possible line of future work (future). Jurgens et al. (2018) then assembled an extensive list of features which could indicate which of these uses a citation is exemplifying, including structural information about where it occurs in the paper, similarity between the cited and citer paper via their abstract, a topic model over sentences which encodes which words occur most commonly together, grammatical constructions which occur commonly (building on the selectional preferences of Erk, 2007), the difference in years between the articles, the venue in which the cited and citer were published, and about 25 others. The authors use 3,083 hand-coded citation contexts along with their functions (from Teufel et al. 2006) to train the model, and use a Random Forest classifier, which works by generating a large set of decision trees for classifying according to these features. When predicting, each set of rules makes their decision, and some average or vote is made from the results of all these decision trees. The results are better than many previous attempts, although still obtain a macro F1 score of only 0.530, indicating that there is still much work to be done in the area.

Teufel et al. (2006) achieve significantly better results with a similar typology, with macro F1 scores closer to 0.7, through a more guided approach. The main difference in their approach was the collection of "cues" by annotators while hand-coding the training set. These cues indicated to the annotator the function of the citation and were then combed by hand by the authors to extract 892 cue phrases, whose presence is then added as an additional feature to the learning algorithm. The clear performance increase through this guided approach is common in natural language processing applications, where contextual knowledge curated by experts can drastically improve a machine learning application in a specific field. Such approaches are often shied away from because of the amount of effort they entail and their lack of generality but should be ideal in my applied context.

Neural networks might be able to understand

The prototypical tool of computational linguistics in the 21st century is the neural network. The neural network is a statistical abstraction inspired by the learning mechanisms of the human brain, and the method has improved the best-in-class in almost every classic field of computational linguistics. Recently neural networks have become incredibly accessible. This has been driven by the ever-lowering cost of computer resources, large open-source projects building intuitive high-level programming libraries for constructing neural networks, and an explosion of work in computational linguistics journals such as ACM, where it is the norm to share data, code, and methods. This all allows researchers to operate neural networks to great profit with minimal technical investment. Neural networks, despite being predominant, are not the only method that is used, but many of the principles of other methods are at a high level the same.

I will present neural networks pragmatically in what follows, neglecting details not relevant for describing the tools needed for this endeavor. A neural network is a weighted directed acyclic network, “neurons” (the nodes) connected along “axons” (the edges). Each node in the network has an *activation*, represented by a single real number. The activation of each node affects all nodes which it connects to. A node’s activation is a function of the weighted sum of the weights of its neighbors, those which have a directed tie to the node. A neural network takes some input (nodes with in-degree zero) and produces some output (nodes with out-degree zero). To train the network, the researcher defines how an output is judged, giving the neurons positive or negative stimulus depending on the quality of their output for given inputs. This judgement (the “loss function”) propagates back through the neurons, changing their connectivity in response to the good or bad outcome. This training by judgement is called *supervised learning*, and through this process neural networks are capable of learning complex and nonlinear patterns from hand-labeled data. Analysis using neural networks shares many best practices with other machine learning methods, such as constructing balanced training examples, choosing the input features wisely, bootstrapping training examples, and employing cross-validation.

The distributional hypothesis and meaning representation

One of the trickiest problems in computational linguistics is how to deal with meaning. The problem is far from solved, and rather than trying to solve it here I will sketch some useful methods for getting at meaning computationally. The most typical methods are distributional, based on the frequency of occurrence of each combination of words. If two words commonly occur close by, there is likely some semantic correspondence between them, such as *dog* and *bone*. The semantic correspondence is that dogs like to chew on bones. Likewise, if two words are typically in the same contexts, they likely have similar meanings. This is evident for occupations such as *optometrist* and *eye doctor*, versus *lawyer*. The first two are exchangeable, and will relate semantically to the same words, e.g. *tooth*, whereas the last would look very strange in those contexts. The proposition that co-occurrence relates meaningfully to meaning is commonly referred to as the distributional hypothesis (Harris 1954).

There has been a surge in methods for distributional semantic analysis on a massive scale, but by far the most prominent approach in recent computational linguistics literature is word2vec (Mikolov et al. 2013). Much of what follows applies to other word embeddings, such as LSA (Deerwester et al. 1990) and the related GloVe vectors of Pennington, Socher, and Manning (2014). Word2vec is a word embedding, meaning it assigns to each word in a vocabulary a list of numbers, a vector. The numbers may or may not have semantic meaning on their own, but what is important is that words which are semantically similar will have vectors which are close to each other. We will see that the vectors assigned by word2vec have another great property, that they can solve analogies algebraically. Word2vec can work in one of two similar ways. In the first, called continuous bag of words (CBOW), a simple neural network is trained to predict a word, given as input a few words directly neighboring as input. The second version of word2vec is called skip-gram and has been found to be most effective in practice. In this method the neural network is given a randomly selected neighboring word and is trained to predict the focal word. The neural network has a single hidden layer, and it is the weights which connect each word to this hidden layer which comprise the N-dimensional vector which represents this word semantically.

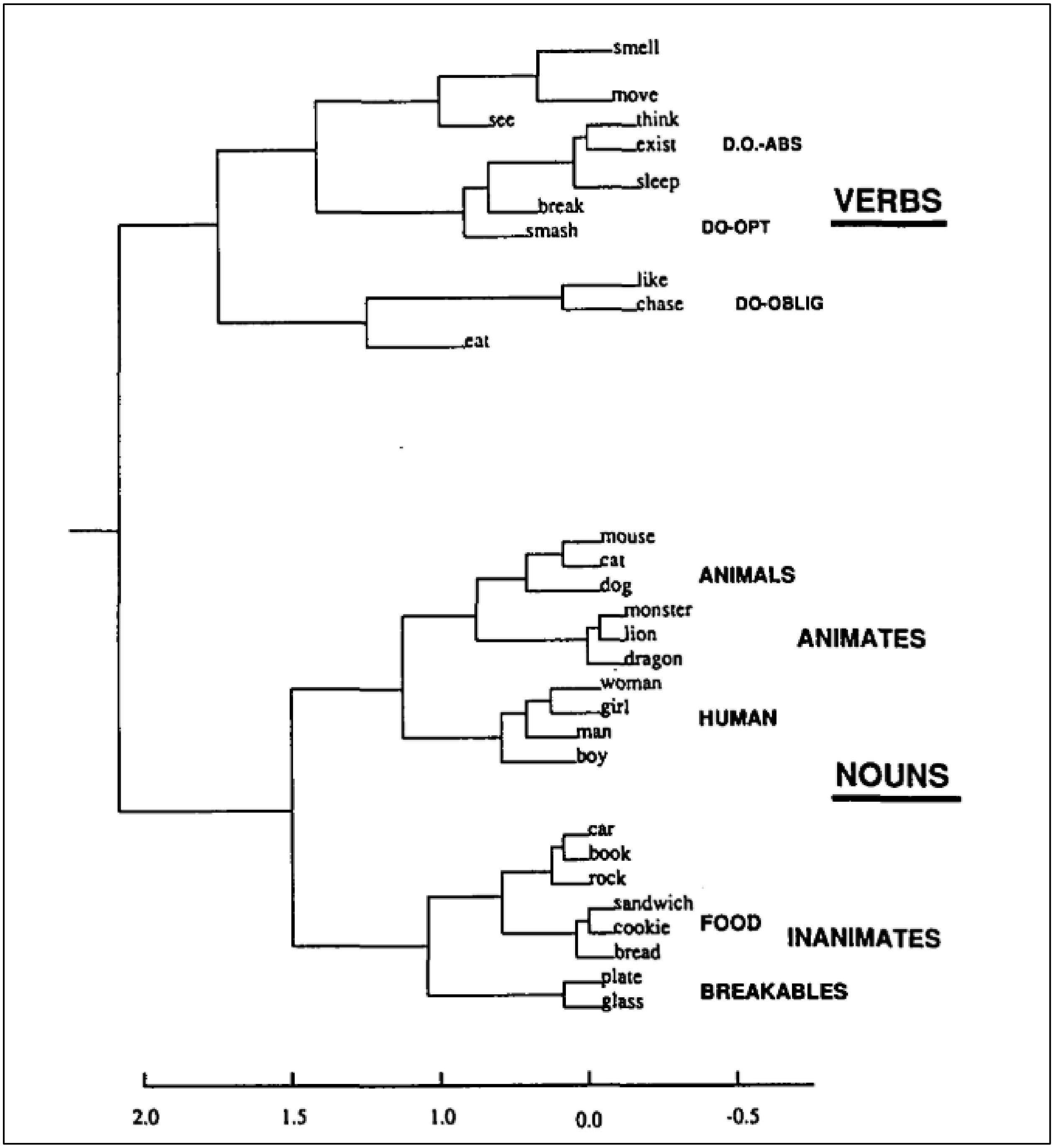

Figure 3. Originally from Elman (1990), this diagram shows the dendrogram from a hierarchical clustering of the vector representations of 31 words used in a simple sentence prediction task. This was one of the first realizations that these internal representations were encoding higher level meaning.

Figure 3 is taken from the first discovery of this property by Elman (1990) and illustrates this property nicely, showing that the vectors for words which are close to each other are semantically similar. The main difference between word2vec and LSA is that LSA considers co-occurrence across the whole document, while word2vec considers only short-range co-occurrences. Also, word2vec is able to train on a much larger set of data (Mikolov et al. 2013) and more adaptable to new notions of meaning besides prediction of nearby words. The remarkable and oft-repeated observation of Mikolov et al. which also makes word2vec special relative to other distributional methods is that this vector representation can solve analogies, most famously:

$$Vec(King) - Vec(Man) + Vec(Woman) = Vec(Queen)$$

where Vec(King) is the N-dimensional vector representation of the word “King”. The authors also showed the superiority of these representations in other semantic natural language processing tasks, where we give a neural network the vector representation as input instead of having an input neuron for each individual word. The method is also quite amenable to extension and modification. For instance, the researcher can change the unit of analysis from words to sentences, paragraphs, or entire documents (Le and Mikolov 2014; Pagliardini, Gupta, and Jaggi 2018). The method should have some leverage in any context where the distributional hypothesis holds (even genetic code, as shown by Asgari and Mofrad 2015). When generating the distributed representations, the research can also modify the task the neural network is trained to do, or even have it complete multiple tasks at once, changing slightly the notion of “meaning” captured by the vectors.

In the context of 100 years of academic literature written from within multiple intellectual communities, word2vec serves the extremely useful purpose of a data-driven thesaurus, identifying which words or multi-word phrases can be used interchangeably in the literature, and when, where, or by whom. This thesaurus allows for a simplification, a canonicalization, of sentences into a common form, permitting me to more easily group together higher-level statements which have the same meaning.

Another fantastic application of word2vec is to track semantic changes over time in large corpora (De Bolla et al. 2019; Hamilton, Leskovec, and Jurafsky 2016; Rodda, Senaldi, and Lenci 2016; Yao et al. 2018). This is done by computing the vector representations of words using temporal slices of the corpus (e.g. just documents from 2011, 2012, and 2013 to produce separate embeddings). These slices will be largely incommensurable with each other, due to the inherently stochastic nature of training neural networks, and the rotation and translation invariance of such embeddings. Researchers compare either by aligning these representations to each other through distance-preserving manipulations, or by tracking the semantic neighborhood of a single word through time. A newer approach is to learn the embeddings at once, using time as a covariate in a sense. See Kutuzov et al. (2018, section 3.2) for an exhaustive review of recent developments in tracking meaning over time, so-called distributional approaches to diachronic semantics.

These basic distributional methods are omnipresent because they perform well in a wide variety of tasks and need only a large corpus of text for training. In representing meaning, however, they are rather crude. For instance they neglect syntax altogether, perhaps encoding that words are related to each other, but not so much how. Also, if single word has multiple meanings, word2vec takes something like an average of these distinct meanings. One way to address these qualms is by analyzing meaning via a word's position in a co-occurrence network. The co-occurrence network draws a link between two words to the extent that they occur in close proximity. We can tune this network, changing how close we ask words to be and the minimum frequency of the co-occurrences for a link to be present. The distributional hypothesis can then be translated into network terminology, that words which are structurally isomorphic in this co-occurrence network have the same meaning.

Exemplifying this method, Small (1980) constructed a co-occurrence network (what he calls a "conceptual map") of core concepts in theoretical chemistry. He derived the list of concepts inductively by looking at the context of citations to the most highly cited works in the field. He then built the network as described above, by drawing a link between two terms to the extent that they are mentioned in close proximity to each other in a single paper. Having identified a conceptual structure in this way, Small then drills down on bridging links, and examines qualitatively the exact nature of this conceptual link by simply reading how the authors relate the two concepts. The papers which span conceptual communities in his conceptual map turn out to be innovative conceptual bridges when qualitatively examined. A useful extension to this, which indeed Small conducted, is to label the edges drawn with the semantic relationship in use. Small worked qualitatively, reading each co-reference, but one could also classify these links based on other words in their common context or the grammatical connectives used.

A conceptual co-occurrence network keeps all distributional information and is more informative than word embeddings, but is much less efficient, and less useful as an input to other machine learning models. Thus this method is ideal for analyzing a small collection of concepts, instead of mapping the meaning of every word in a language. Hamilton et al. (2016) show an additional use of this network to measure how polysemic a word is, how many meanings it can obtain, by its local clustering coefficient in the conceptual network (Watts and Strogatz 1998). If a word has multiple distinct meanings, it should show up in multiple distinct contexts, and thus be connected to more than one cluster of co-occurring words. This multiple-membership information is completely unavailable in any vector embeddings of words, which implicitly assume there to be no polysemy. Identifying and tracking polysemy in the usage of theoretical terms over time bears directly on hypotheses that such multifunctionality is instrumental for the long life of theories and theoretical concepts. It also potentially allows me to identify when, where, and how academic sociology reached consensus, and if at all.

Syntactic and linguistic nuance

One strong limitation of distributional methods is their total ignorance of the content of these semantic relationships. Work in computational linguistics is at its most useful in providing methods to represent and extract the grammatical structure of sentences. The most widespread grammatical representation is the dependency grammar. A dependency grammar decomposes a sentence into a labeled tree of its composite words, with the main verb as the root. Each word in the sentence is connected to the main verb either directly or indirectly through labeled edges. For instance, the subject of the sentence is connected via an edge labeled “nsubj” and nouns which are part of a conjunction are connected via “conj”. Figures 1a and 1b show the result of SpaCy’s dependency parsing algorithm on the sentence “In this comment I will first review Parson’s hypothesized four stages in Durkheim’s theoretical development.” (Pope 1975) For an extended overview of dependency parsing, see Jurafsky and Martin (2000, Ch. 13). This representation makes the grammatical structure of a sentence available to computational analysis and is useful because of the massive amount of work that has gone into creating effective parsers. The Stanford NLP Group provides multiple highly optimized state-of-the-art and pre-trained parsers for free use (e.g. Chen and Manning 2014). Other alternatives exist, such as the open-source SpaCy, Google’s Parsey McParseface, and the Berkeley parser, but their differences in accuracy are miniscule in the meticulously structured discourse of academic text.

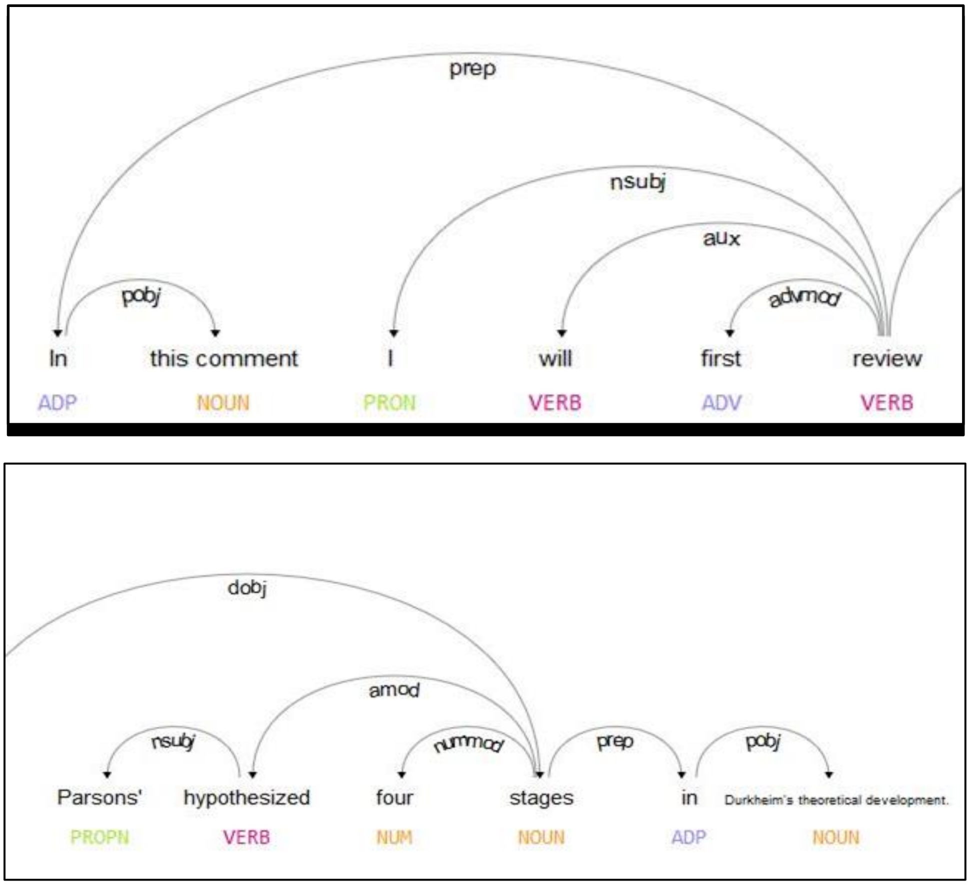

Figure 1a / 1b: Example dependency parse by SpaCy of a randomly selected sentence about Durkheim. Visualization was produced using displaCy at https://explosion.ai/demos/displacy. The figure has been cut in half for readability.

One potentially useful strategy for tracking the understandings of concepts through our dataset is in a syntactic analysis of how they are connected to other words and concepts. This strategy is inspired by work in critical discourse analysis (as explained to me in Hart 2014), a field which has great sensitivity for what kinds of verbs are applied to which discursive actors. For instance, Hart (2014:24) analyzes articles from two ideologically divergent newspapers concerning protestors' interactions with police. In one newspaper the protestors are typically the subject of strong and violent verbs ("caused," "dragged," or "broke") whereas in the other newspaper the policemen are the subjects of similarly violent words ("provoke," "split," "moved into"). A syntactic analysis of academic discourse can most simply identify what authors are allowed to do with their research. Bertin et al. (2016) takes this approach using a simple word frequency analysis, looking at the verbs which appear in citation contexts across PLOS. What shows up are phrases such as “has been found”, “it is known”, “has demonstrated”, “has proposed”, “previously reported”. The authors find differences in agency at different points in the structure of the article.

The closest that researchers have gotten to encoding abstract meaning, the who does what to what, at what time, in which place, in a machine-readable format, is through meaning representation formats, most notably Abstract Meaning Representation (AMR). AMR, like dependency grammar, represents meaning of a sentence as a directed labeled tree. The difference is where dependency grammar uses surface representations (words) as nodes, AMR resolves these surface representations to concepts. Verbs have arguments in AMR, which must be filled precisely by the other concepts present in the text. AMR parsing has not made quite as much progress as dependency parsing, however, and has only been used effectively on rather simplistic sentences. Gathering training data, moreover, is highly labor intensive, and modifications would likely have to be made to training algorithms for academic statements, a highly technical task. Thus although AMR is an ideal representational format for dealing with the meaning contained in academic texts, it will likely have to wait at least another decade before being capable of out-of-the box application to my context. That being said, AMR gives some inspiration for what the meaning of an academic text when encoded in a machine-readable way would look like, and indeed researchers in computational linguistics have shown us some of the potential of such a dataset, constructing algorithms which can answer questions, search intelligently, or identify contradictions or logical entailments.

Some simpler tools

In this section I present some of the simpler tools available for a machine-assisted analysis of a large body of literature. These simple tools are often most effective, as the researcher and the hypothetical reader of a results section has no trouble understanding what has been done and why it is useful.

First, to direct qualitative analyses of all literature published in these journals, I have indexed the documents into an open-source search engine, Lucene[2]. This application provides me with immediate access to every mention of a word or group of words, to any author, etc., and allows me to very quickly construct simple counts and to examine the context of these mentions.

One puzzle which presents itself to many attempts to understand computationally, or even to search effectively, is how we can identify when two sentences, or more simply two words, have the same or similar meaning. A crucial tool in this effort is WordNet (Miller 1995). WordNet is a dictionary and thesaurus accessible in Python. The dictionary is built of "synsets", groupings of words which have the same meaning. It also features hypernym-hyponym relationships between synsets, indicating where one concept is contained within another, as {bunkbed} is to {bed}. In a similar manner, adjectives are connected to their opposites, verbs are connected to related, more extreme, or logically entailed verbs. Each word in a synset also includes a database of different ways of expressing it (a word's "lemmas"), including for example its plural. This conceptual network allows me to directly import commonsense logical relations between words without much worry for completeness or correctness, as the dictionary has been curated by hand. This dictionary also gives some of the technical leverage to disambiguate between different senses of a word in context, by informing us that there are such senses and their semantic relationship to other words. The primary limitation of this dictionary, addressed below by distributional semantic methods, is that it is limited to words in common English, and will have no idea for example what "role set" refers to, or what synonyms it might have.

In a similar vein, Wikidata allows me to import an incredible number of semantic relationships into my analysis, but its generality trades off with its completeness and correctness. Like WordNet, Wikidata is a graph database of relations between concepts, although because of its size it is stored in Wikidata servers and must be queried. Unlike WordNet, Wikidata includes people, places, institutions, journals, and a plethora of more context-specific conceptual entities. For instance, the entry for "Max Weber" contains 136 "statements," from the simple fact that he is a

[2] Apache Lucene is a collection of open-source search software. More information can be found online at https://lucene.apache.org/

human to his country of citizenship, place and date of birth, his various employers, doctoral advisors, etc. (see https://www.wikidata.org/wiki/Q9387 for the full entry). Wikidata is useful for instance in constructing small lexicons, for instance of the name of every political party ever existing in Germany, a comprehensive set of adverbs, a list of 21st century technologies, etc. Because the database is queried via what is called SPARQL, we can ask more complex questions such as "What scientists who were born between 1927 and 1937 have a son named John who lives in Delaware?"

Conclusion

A published academic work is an amalgam of a complex process of socializing into a community of discourse and of acceptable research activity, of painstakingly writing and editing, possibly collaboratively, and of responding to reviewers. It is a presentation, an attempt at achieving some academic recognition, or at least at being published at all. It is justification that you are actually an academic, doing academic work, and for instance that you should receive a job, or tenure. It is a genuine attempt at constructing new knowledge, or at communicating knowledge already attained, whatever your philosophy may be. It is all of these things differentially, an exceptionally nuanced social activity. Even as I constructed this report I find myself struggling in the endeavor which has produced my object of study. And in sociological publications there is reserved a very special place for classical authors and ideas. It is uncommon that papers will directly test the theoretical wisdom of the classics (see Adatto et al. 1981), but they are used constantly in justifying the research at hand. Indeed, it is not so important with what argument or data Weber, Durkheim, Simmel, or even minor but trusted figures use to arrive at an understanding of the social world, only that they wrote it down somewhere which can be cited. This very week a mentor of mine advised me not to look too hard for prior understandings of a concept he had coined, unless it was a revered classical name which might legitimize the concept. And yet the classic works for the most part collect dust on the shelves while researchers collect and use second- and third- hand interpretations and essentializations or crack the tome to read just those statements relevant to their justification practices.

Of course this is a partial truth based on my own preconceptions and biases but begs a deeper look into these interpretations. Collecting, categorizing, relating and dissecting these interpretations, in short cataloguing them, gives us some foothold to trace patterns in this interpretive process, building the raw materials for a theory of the development and maintenance of this system of knowledge. Ideally it would not just confirm my precocious cynicism but would produce a more refined and delimited understanding of the development of these understandings through time. And even if this is exceedingly optimistic, I will have dug a little deeper than most into the literature of our field, produced a bit more understanding for those around me, and learned something about bibliometric analysis and computational linguistics via a project I can get obsessed about.

Appendix A – The nitty gritty of processing 20k academic papers

I have relegated to this appendix some of the most underwhelming technical details which need to be accounted for in my bibliometric and computational linguistic analysis of sociological literature, removing them from the body of this essay where they seemed an unneeded distraction.

First, because I have been provided with papers in a plaintext format, I need to manually parse the bibliographies and in-text citations, clean headers and footers of the document and split the document into sections, and rely on tools developed by computational linguists for this task. The problem of identifying which papers are referenced in any given paper, and where, using the full text of the bibliography has essentially been solved by researchers in computational linguistics, specifically in ParsCit (Councill, Giles, and Kan 2008) and its successor Neural ParsCit (Prasad, Kaur, and Kan 2018). Both methods function by labeling words in a bibliography with which part of the information in a citation they refer to, e.g. author, year, or journal. Both approaches are entirely open source and permit me to re-train their models myself. ParsCit handles the problem using the conditional random field statistical abstraction (Lafferty et al., 2001). As the name would imply, Neural ParsCit solves the problem using neural networks, and presents a statistically significant, albeit small, improvement on ParsCit. In-text citations, being quite uniform in structure, succumb easily to a more explicit non-statistical strategy (see e.g., Prasad et al. 2018). Following their lead, I have constructed regular expressions to parse them. Regular expressions give explicit deterministic parsing rules, specifying exactly what the surface form should look like. For example, a regular expression which parses a list of authors within a citation would look like “{authorName}, {authorName}, and {authorName}”, or just “{authorName} et al.”, where {authorName} is another regular expression which matches a capital letter with some number of upper or lowercase letters following[3].

Another important task boring enough to appear in this appendix is how to turn a book which has not already been digitized into machine-readable text. Optical character recognition (OCR) is the main component in this process, and takes as input an image, for example of a page of a book, and returns the text on the page. After much experimentation I have settled on tesseract, an open source OCR software built by Google, which besides being scary accurate and relatively efficient has the useful additional feature of giving the coordinates of a box which contains each word it detects. The algorithm also detects lines (e.g. the line in this paragraph which contains *this* word) and deals intelligently with multiple column format. This output must then be post-processed to remove headers, footers, and page numbers, which is trivial as long as they occur in the same place on every page of a book.

One small operation I have to conduct on all my text is to reverse the hyphenation breaks across lines. A hyphenated word *hyphenation* will show up in my dataset as “hyph-(line break) enation”. Fortunately, this is an atypical pattern, and a simple RegEx replacement does very well, replacing “([a-z])-\s*\n\s*([a-z])” with “\1\2”, the two letters which have been split by the dash.

I will also likely have to go through the process of re-training existing methods or training my own machine learning algorithms on my dataset, as there are differences between computer science and sociological references, between linguists’ use of citations and those of sociologists of the 1930s. As this example hints, my dataset spans such a long time that styles even of references could be variable within my dataset. The crucial tool for training a neural tagger, for instance, is an interface for human tagging, for which I plan to use the brat rapid annotation tool

[3] These are expressed in RegEx as:

```
    name = "[A-Z][A-Za-z-]+"
    fname = "({name}(\s{name})?|{name} et al.?)"
    names = "{fname}(, {fname})*(,? (and|&) {fname})?"
```

(brat.nlplab.org) which has incorporated both word-range tagging and acyclic labeled graph tagging of various kinds, allowing for dependency tagging or even AMR tagging.

Bibliography

Adatto, Kiku and Stephen Cole. 1981. “The Functions of Classical Theory in Contemporary Sociological Research: The Case of Max Weber.” *Knowledge and Society: Studies in the Sociology of Culture* 3:137–62.

Asgari, Ehsaneddin and Mohammad R. K. Mofrad. 2015. “Continuous Distributed Representation of Biological Sequences for Deep Proteomics and Genomics.” *PLoS ONE* 10(11):1–15.

Austin, John Langshaw. 1975. *How to Do Things with Words*. Oxford University Press.

Baehr, Peter. 2016. *Founders, Classics, Canons: Modern Disputes over the Origins and Appraisal of Sociology’s Heritage*. Second. New Brunswick: Transaction Publishers.

Belter, Christopher W. 2016. “Citation Analysis as a Literature Search Method for Systematic Reviews.” *Journal of the Association for Information Science and Technology* 67(11):2766–77.

Blumer, Herbert. 1969. *Symbolic Interactionism: Perspective and Method*. Berkeley: University of California Press.

De Bolla, Peter, Ewan Jones, Paul Nulty, Gabriel Recchia, and John Regan. 2019. “Distributional Concept Analysis.” *Contributions to the History of Concepts* 14(1):66–92.

Breiger, Ronald L. 1974. “The Duality of Persons and Groups.” *Social Forces* 53(2):181–90.

Chen, Chaoemi. 2006. “CiteSpace II: Detecting and Visualizing Emerging Trends and Transient Patterns in Scientific Literature.” *Journal of the American Society for Information Science and Techonology* 57(3):359–77.

Chen, Chaomei, Fidelia Ibekwe-SanJuan, and Jianhua Hou. 2010. “The Structure and Dynamics of Cocitation Clusters: A Multiple-Perspective Cocitation Analysis.” *Journal of the American Society for Information Science and Technology* 61(7):1386–1409.

Collins, Randall. 1998. *The Sociology of Philosophies: A Global Theory of Intellectual Change*. Cambridge: Harvard University Press.

Corominas-Murtra, Bernat, Carlos Rodríguez-Caso, Joaquín Goñi, and Ricard Solé. 2010. “Topological Reversibility and Causality in Feed-Forward Networks.” *New Journal of Physics* 12.

Davis, Murray S. 1986. “‘That’s Classic!’ The Phenomenology and Rhetoric of Successful Social Theories.” *Philosophy of the Social Sciences* 16(3):285–301.

Deerwester, Scott, Richard Harshman, Susan T. Dumais, George W. Furnas, and Thomas K. Landauer. 1990. “Indexing by Latent Semantic Analysis.” *Journal Of The American Society For Information Science* 41(6):391–407.

Ding, Ying, Guo Zhang, Tamy Chambers, Min Song, Xiaolong Wang, and Chengxiang Zhai. 2014. “Content-Based Citation Analysis: The Next Generation of Citation Analysis.” *Journal of the Association for Information Science and Technology* 65(9):1820–33.

Doreian, Patrick, Vladimir Batagelj, and Anuska Ferligoj. 2004. *Generalized Blockmodeling*. Cambridge: Cambridge University Press.

Elman, Jeffrey L. 1990. “Finding Structure in Time.” *Cognitive Science* 14:179–211.

Erk, Katrin. 2007. “A simple, similarity-based model for selectional preferences.” Pp. 216–223 In *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics (ACL)*.

Fine, Gary Alan and Sherryl Kleinman. 1986. “Interpreting the Sociological Classics: Can There Be a ‘True’ Meaning of Mead?” *Symbolic Interaction* 9(1):129–46.

Garfield, Eugene, Irving H. Sher, and Richard J. Torpie. 1964. *The Use of Citation Data in Writing the History of Science*. Philadelphia: Institute for Scientific Information.

Geertz, Clifford. 1973. *The Interpretation of Cultures*. New York: Basic Books, Inc.

Hamilton, William L., Jure Leskovec, and Dan Jurafsky. 2016. “Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change.” *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* 1489–1501.

Harris, Zellig S. 1954. “Distributional Structure.” *WORD* 10(2–3):146–62.

Hummon, Norman P. and Kathleen Carley. 1993. “Social Networks as Normal Science.” *Social Networks* 15(1):71–106.

Hummon, Norman P. and Patrick Doreian. 1989. “Connectivity in a Citation Network: The Development of DNA Theory.” *Social Networks* 11:39–63.

Jurgens, David, Srijan Kumar, Raine Hoover, Dan McFarland, and Dan Jurafsky. 2018. “Measuring the Evolution of a Scientific Field through Citation Frames.” *Transactions of the Association for Computational Linguistics* 6:391–406.

Jurafsky, Daniel and James H. Martin. 2000. *Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition.* Upper Saddle River, N.J:Prentice Hall.

Kuhn, Tobias, Matjaž Perc, and Dirk Helbing. 2014. “Inheritance Patterns in Citation Networks Reveal Scientific Memes.” *Physical Review X* 4(4):1–9.

Kutuzov, Andrey, Lilja Øvrelid, Terrence Szymanski, and Erik Velldal. 2018. “Diachronic Word Embeddings and Semantic Shifts: A Survey.” Pp. 1384–97 in *Proceedings of the 27th International Conference on Computational Linguistics*.

Le, Quoc and Tomas Mikolov. 2014. “Distributed Representations of Sentences and Documents.” *Proceedings of the 31st International Conference on Machine Learning, PMLR* 32(2):1188–96.

Lin, Chi Shiou. 2018. “An Analysis of Citation Functions in the Humanities and Social Sciences Research from the Perspective of Problematic Citation Analysis Assumptions.” *Scientometrics* 116(2):797–813.

Liu, John S., Louis Y. Y. Lu, and Mei Hsiu-Ching Ho. 2019. “A Few Notes on Main Path Analysis.” *Scientometrics* 119(1):379–91.

Liñán, Francisco and Alain Fayolle. 2015. “A Systematic Literature Review on Entrepreneurial Intentions: Citation, Thematic Analyses, and Research Agenda.” *International Entrepreneurship and Management Journal* 11(4):907–33.

Malliaros, Fragkiskos D. and Michalis Vazirgiannis. 2013. “Clustering and Community Detection in Directed Networks: A Survey.” *Physics Reports* 533(4):95–142.

Masterman, Margaret. 1970. “The Nature of a Paradigm.” Pp. 59–89 in *Criticism and the Growth of Knowledge*, edited by Imre Lakatos and A. Musgrave. Cambridge: Cambridge University Press.

McMahan, Peter and James Evans. 2018. “Ambiguity and Engagement.” *American Journal of Sociology* 124(3):860–912.

Miller, George A. 1995. “WordNet: A Lexical Database for English.” *Communications of the ACM* 38(11): 39-41.

Moravcsik, Michael J. and Poovanalingam Murugesan. 1975. “Some Results of the Function and Quality of Citations.” *Social Studies of Science* 5:86–92.

Mouzelis, Nicos. 1995. *Sociological Theory: What Went Wrong? Diagnoses and Remedies.* London: Routledge.

Newman, M. E. J. and M. Girvan. 2004. “Finding and Evaluating Community Structure in Networks.” *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 69(2 2):1–15.

Pagliardini, Matteo, Prakhar Gupta, and Martin Jaggi. 2018. “Unsupervised Learning of Sentence Embeddings Using Compositional N-Gram Features.” 528–40.

Parsons, Talcott. 2010. *The Structure of Social Action*. New York: The Free Press.

Pennington, Jeffrey, Richard Socher, and Christopher Manning. 2014. “Glove: Global Vectors for Word Representation.” Pp. 1532–1543 in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*.

Perc, Matjaž. 2013. “Self-Organization of Progress across the Century of Physics.” *Scientific Reports* 3:1–5.

Rodda, Martina A., Marco S. G. Senaldi, and Alessandro Lenci. 2016. “Panta Rei: Tracking Semantic Change with Distributional Semantics in Ancient Greek.” *CEUR Workshop Proceedings* 1749.

Rosvall, M. and C. T. Bergstrom. 2008. “Maps of Random Walks on Complex Networks Reveal Community Structure.” *Proceedings of the National Academy of Sciences* 105(4):1118–23.

Shwed, Uri and Peter S. Bearman. 2010. “The Temporal Structure of Scientific Consensus Formation.” *American Sociological Review* 75(6):817–40.

Small, Henry. 1980. “Co-Citation Context Analysis and the Structure of Paradigms.” *Journal of Documentation* 36(3):183–96.

Teufel, Simone. 1999. *Argumentative Zoning: Information Extraction from Scientific Text*. Ph.D. dissertation. University of Edinburgh.

Teufel, Simone, Advaith Siddharthan, and Dan Tidhar. 2006. “Automatic classification of citation function.” Pp. 103-110 in *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP).*

Teufel, Simone, Advaith Siddharthan, and Dan Tidhar. 2010. “Automatic Classification of Citation Function.” (July):103.

Volanakis, Adam and Konrad Krawczyk. 2018. “SciRide Finder: A Citation-Based Paradigm in Biomedical Literature Search.” *Scientific Reports* 8(1):1–7.

Watts, D. J. J. and S. H. H. Strogatz. 1998. “Collective Dynamics of ‘small-World’ Networks.” *Nature* 393(6684):440–42.

White, Howard D. 2004. “Citation Analysis and Discourse Analysis Revisited.” *Applied Linguistics* 25(1):89-116+132.

Xie, Zheng, Zhenzheng Ouyang, Pengyuan Zhang, Dongyun Yi, and Dexing Kong. 2015. “Modeling the Citation Network by Network Cosmology.” *PLoS ONE* 10(3):1–13.

Yao, Zijun, Yifan Sun, Weicong Ding, Nikhil Rao, and Hui Xiong. 2018. “Dynamic Word Embeddings for Evolving Semantic Discovery.” 673–81.

Zhou, Yuan, Fang Dong, Dejing Kong, and Yufei Liu. 2019. “Unfolding the Convergence Process of Scientific Knowledge for the Early Identification of Emerging Technologies.” *Technological Forecasting and Social Change* 144(May):205–20.