Knowledge Graphs and Network Text Analysis

Popping, Roel

Published in:
Social Science Information

DOI:
10.1177/0539018403042001798

Social Science Information

http://ssi.sagepub.com

Knowledge Graphs and Network Text Analysis

Roel Popping

Social Science Information 2003; 42; 91

DOI: 10.1177/0539018403042001798

Roel Popping

Knowledge graphs and network text analysis

Abstract. *A knowledge graph is a kind of semantic network representing some scientific theory. The article describes the present state of this field and addresses a number of problems that have not yet been solved. These problems are implicit relations, strength of (causal) relations, and exclusiveness. Concepts might be too broad or complex to be used properly, so directions for solving these problems are explored. The solutions are applied to a knowledge graph in the field of labour markets.*

Key words. *Knowledge graphs – Knowledge representation*

Introduction

One of the fields in which networks are used is text analysis. The networks are usually applied for evaluation purposes or to compare (cognitive) maps. They might also represent a state of the art in some field. All kinds of techniques are available to make these evaluations, comparisons or representations more manageable. In this article the representation of knowledge is considered. Insights from the evaluation and comparison approach are used.

Explicit knowledge can be represented in many ways. Explicit knowledge is knowledge that has been formulated. Recited and written texts are among the oldest ways of representation. These can be clear, but also obscure. In general the full knowledge or problem area is not presented, so that the listener or reader does

Social Science Information © 2003 SAGE Publications (London, Thousand Oaks, CA and New Delhi), 42(1), pp. 91–106.
0539-0184[200303]42:1;91–106;031798

not immediately get a complete overview. To get such an overview, knowledge can be structured in schemes. These might consist of hierarchical tree structures (assuming the knowledge allows such) or arrow diagrams, which indicate what is or might be influenced by what. The ambiguity of language is a problem; concepts are often interpreted in several ways (Popping, 2000). A second problem is that knowledge is dynamic, while text is a static means of knowledge representation. A change in knowledge implies that parts of the text have to be reworked.

Approaches to knowledge representation are classified into two major groups (Rich, 1983). The first group is logical formalism. This includes, among others, predicate logic. Semantic networks belong to the other group, structured representation. Logical formalism is found in expert systems. An expert system is a computer application that performs as if it were a human expert. It uses a rule base, a knowledge base, and an inference engine. It has an input and an output system. The rule base contains procedural knowledge (Stokman and De Vries, 1988: 188), which consists of a set of prescriptions for actions and is often presented as a set of if–then rules. The knowledge base contains a declarative representation of knowledge and is given as a set of assertions about a certain subject. Conclusions can be drawn from these assertions by inference methods. The goal of a semantic network is to make clear the structure of knowledge, so it aims at a structured representation.

One specific kind of semantic network is the so-called knowledge graph. Such graphs use only a few types of relations. In addition they allow new knowledge to be added to the graph. This knowledge can be integrated with the already available knowledge. In these graphs procedural knowledge is thus used for the verification of rules. Conceptual knowledge is used for the integration of definitions and causal models. Some rules regarding the transformation of text into graphs exist.

We are interested specifically in the representation of sociological knowledge, especially in knowledge that results from empirical research. Now some issues are particularly important. One issue concerns exclusiveness. Knowledge is not always general, but holds under certain conditions, and might be exclusive for a certain group. Empirically proven sociological knowledge might be found in a study among women. The original investigator has not concentrated on men, or has not been able to prove the findings for men. This has repercussions on the representation of the knowledge. The knowledge is exclusive for women. Another issue concerns the choice of concepts. What is a useful concept? In a network such a concept is represented as a point. Often a concept includes a process in which several "smaller" concepts are related. Sometimes this is necessary, but certainly not always. For this investigation, we concentrate on knowledge in the area of labour markets.

The article is structured as follows. First the notion of knowledge graphs is introduced. Next the issues mentioned above are introduced and elaborated. These extensions are applied to a knowledge graph on theories about labour markets. A discussion follows on what can be said by using this graph. Finally conclusions are drawn.

Knowledge graphs

The representation of knowledge by labelled nodes and links between these nodes leads to structures that are usually called semantic networks (Sowa, 1987). Knowledge graphs can be viewed as a particular kind of semantic network. One essential difference between knowledge graphs and semantic networks is the explicit choice of only a few types of relations (James, 1992: 98).

The construction of a knowledge graph starts with the extraction of information from texts. This is called text analysis. The result is a list of concepts, represented as labelled points, and a list of typed links between the points. These form the so-called author graph.[1] A concept is a unit of meaning. It is used as the basic unit for the meaning content of what it refers to (Popping, 2000: 17). The most important type of link between points is the causal relation. The goal of the group of scientists working on knowledge graphs was to construct graphs that represent the theory on a specific subject. So far this is identical with the methods that have been discussed.

The next step is called concept identification. Here the various author graphs are combined into one graph by identifying points with each other. When the texts that were the basis of the graphs deal with the same subject, points with the same label are identified. An author may use synonyms for a concept; therefore points with different labels should be identified. This is done by comparing the neighbourhoods of points to identify the potentially identical pairs.

An index has been developed for measuring the similarity between two sets of points. The value this index takes, in combination with a threshold value, can be used to decide upon identification of two concepts. In the same way it is possible to detect points with the same label, but referring to a different content, so-called homonyms. For example, a chair is something one sits on, but it can also refer to one's position (e.g. a committee leader or a professor). One of these points should receive another label. A compiled graph results which is free of ambiguity of language. This graph is further investigated in procedures called concept integration and link integration. The first procedure tries to find interesting substructures; the second procedure infers new links from the given ones.

The result is called the integrated graph. In order to represent the structure of knowledge, a complex relation is often necessary: the frame relation. This relation combines a number of concepts and relations that are inseparably connected into a single concept. These concepts and relations together ensure that the frame functions as it is supposed to. An example of a frame is the measurableness of quality of work. A bicycle might also be regarded as a frame; it consists of the frame, the wheels, the handlebar, and so on. Together all these parts enable the bicycle to work. Concept integration aims at determining those subgraphs that are candidates for contraction into a frame. (Note that the term is used in a different sense from that generally used in artificial intelligence.) In link integration relations are combined to deduce new relations. If there exist relations between points A and B as well as between B and C, there may be reasons to infer a relation between A and C. To find these new relations, path algebra (Carré, 1979: 84–5) is used. Relations can be based on multiplication, for a serial combination, and addition, for a parallel combination.

Four characteristics are distinguished with respect to relations: directionality, meaning, sign and strength (Carley, 1993; Popping, 2000: 99). In the theory of knowledge graphs, the first two have been used so far. All relations are unidirectional, and the meaning is denoted by using types (see below). A relation like "is friends with" is not used, but "is a kind of" *is* used. The characteristics sign (positive or negative) and strength (usually a value on a 0–1 scale) have up to now not been used.

Originally the idea was to represent knowledge by using as few semantic relations as possible. First only the types CAU, PAR and AKO were used. CAU denotes a cause–effect relation (unstable market positions cause polarization.). The relation is asymmetric and transitive. In all methods using networks based on text, the causal relation is read as ‘‘*might* cause’’. PAR stands for ‘‘is part of’’ (having relations with high status is a part of social capital); it characterizes a thing. AKO refers to ‘‘is a kind of’’ (a married man is a kind of man; a poorly educated woman is a kind of woman); here something is exemplified. The latter two relations are transitive and asymmetric.[2]

Inverse relations are also distinguished, viz. CBY (is caused by), HAK (has as kind) and HAP (has as part) (Stokman and De Vries, 1988). Today different types of concepts are distinguished: tokens and types. Tokens play a role similar to that of variables in logic. Types are labelled points representing generic concepts that are determined by their attribute sets. They can be seen as giving schema information, whereas tokens represent arbitrary instantiations of types. A token denotes an individual that can be chosen from a universe given by the discourse. As an example, ‘‘Pluto’’ is a token, and ‘‘dog’’ is a type. The choice of the individual might be restricted, and the restriction follows from the relations attached to the token. The relation between token and type is denoted by ALI (alike). There are seven relations between types: PAR, CAU, AKO, ORD (ordering), ASS (symmetric association), EQU (equal, symmetric) and DIS (distinct). The label Θ represents a token. Now any arbitrary person can be represented by Θ ALI worker, but the individual John is represented as John EQU Θ ALI worker. This is preferred over John AKO worker, because the AKO relation links an instantiation with an attribute set, and mixes up extension with intension. This is expressed by John PAR workers and John ALI worker. This strict distinction between ‘‘worker’’ and ‘‘workers’’ is not always made in semantic network theory.

Knowledge graphs have, among other things, been applied to theories of mathematical modelling (Bakx et al., 1987), knowledge representation theory (Van den Berg and Hoede, 1991), labour market theories (Popping and Strijker, 1997) and definitions of the concept of imperialism (Hoede and Weening, 1999). Popping (1998) has emphasized the role knowledge graphs might play in process management. Yuen and Richards (1994) have proposed a method for theory construction in qualitative research, which closely resembles the knowledge graphs approach.

Input for knowledge graphs

Popping and Strijker (1997) argued that the texts on which a knowledge graph is based have to contain actual information that has been tested in empirical research;[3] desirabilities and opinions are not allowed. Furthermore it was stated that concepts must be well defined and, in the case of causal relations, the direction of the change must be known. It is not known whether it is better to use complete texts or only parts of texts. The investigators started with the summaries of doctoral dissertations, but these did not contain the essential information; not all the relevant findings were discussed in these summaries, and a great deal of attention was devoted to new questions, etc.

The process of building knowledge graphs starts with extracting information from texts. In re-examining the earlier analyses, four issues come up, for which a solution will be suggested based on developments from network text analysis. The first three issues concern the linking process and the fourth regards the concepts.

First, texts contain implicit knowledge, or there exists implicit knowledge between texts. This is knowledge that is not recognized by a computer, but is highly relevant. We investigated one study about girls, and another one about women. The relation between girls and women belongs to implicit knowledge. Second, so far in knowledge graphs certain relation types are not treated correctly. The causal relation is presented as being deterministic. In research, however, the causal relation is usually expressed as some value, often a regression coefficient. This refers to the preceding remark. Third, some knowledge is often true for only a specific subgroup. The studies above were about women, therefore the findings do not hold for men. At least this was not shown. Hereafter solutions for these problems are suggested based on what is known from network text analysis. The main research groups in this field concentrate on map-comparison analysis (Carley, 1986; Carley and Palmquist, 1992) and on the analysis of evaluative texts (Van Cuilenburg et al., 1986, 1988). Both approaches are summarized by Popping (2000).

The fourth issue concerns the fact that concepts are used as defined by the original investigators. These concepts might resemble each other very closely, but they are defined in different ways. This might make it look as though there are more differences than in reality exist. Instructions for solving this problem will be given later.

Implicit knowledge

As has already been indicated, one of the studies investigated by Popping and Strijker concerned girls who had finished their training and had to look for employment. Another study looked at the position of women. For the reader, it is common knowledge that a girl is a (kind of) woman, and that in general what holds for women is also true for girls. This knowledge is relevant in the process of linking. In the concept identification process, however, this implicit knowledge will not automatically be recognized, nor, as the concepts come from studies that deal with different problems, should one expect beforehand that, on comparing the neighbourhood, it will be found that these are (almost) identical concepts. The investigator may decide that the concept "girl" is identical to the concept "woman", in which case the one should be replaced by the other, but the investigator might also decide that there are here two concepts between which some relation exists. I prefer this option.

In her studies Kathleen Carley (1986, 1988) is also confronted with implicit knowledge. Texts and transcripts are produced within social contexts about which an investigator can obtain expertise. Using such expert knowledge, it becomes possible to "fill in" mental maps with concepts and relations beyond those explicitly stated in the texts under analysis. Carley (1988) has developed an expert system (Social Knowledge Interpreter – SKI) for filling in maps based on an expert's knowledge. The basic idea behind SKI is that, when individuals use certain concepts, other concepts and relations are implied. Much of this implied information is social knowledge shared by speakers (or writers) and audiences (or readers).

A similar tool can be used in the case of knowledge graphs. This can be simply a separate knowledge graph that will be integrated with the other graphs and will include the implicit knowledge. This is, among others, the knowledge that will explain the relation between girl and woman, e.g. girl AKO woman. The knowledge is declarative; it does not contain an interpretation (apart from common sense) as it does in Carley's situation. It is not conceivable to create such a graph in advance. The investigator finds the implicit knowledge when inspecting knowledge graphs, and then makes this knowledge explicit. This allows control by others.

Strength of a relation

The strength of a relation in knowledge graphs has so far not been discussed. This implies that the strength was always set at 1 (one). In evaluative network analysis, a crucial question is how a concept is evaluated by a certain medium (for example, how does the medium "*New York Times*" evaluate the concept "President Bush"?). For this evaluation, the strength of a relation is indicated on a scale from 0 to 1. This makes it possible to distinguish between expressions like "rather good" and "very good". I do not want to evaluate, but to indicate the strength between two concepts as additional information. In the situation of a causal relation between concepts A and B, a strength of 1 would imply that, if A occurs, B will also occur (a sufficient condition). In empirical research, however, a correlation or regression coefficient usually indicates the strength of a causal relation. This might be read as a probability. Here the strength can add a useful meaning. This use of strength applies especially to three types of relation: the cause, the symmetry, and the part of relation.

The correlation or regression coefficient can serve as an indication of the strength of a causal relation. The correlation coefficient merely indicates the association between two concepts. In the case of a regression coefficient, this association is placed within a theoretical model, so the restrictions imposed by the model also apply. The strength of such connections might be incorporated in the knowledge graph. Now the user can decide how to value this strength. Preferable is a standardized association or correlation index. Such an index ranges from –1 to 1, and the value 0 is found in the case of no association.

Strength might also be useful as a heuristic value in the situation of a similarity or "part-of" relation. As an example, the strength might indicate how probable it is that the own-level of educational training is part of the employee (and not of someone else), or how similar two concepts are. If they are perfectly similar, i.e. they are equal, the strength is 1. But "something in common" might also be an indication of "similar". In this case the strength is less than 1. It is preferable to reserve the strength for situations in which causality or association is expressed. In other situations the strength is questionable.

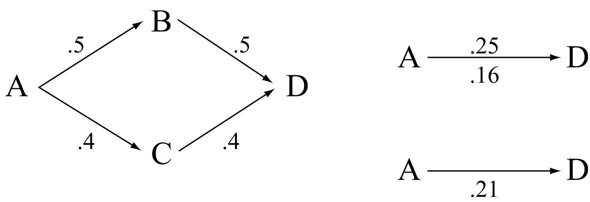

FIGURE 1
Strength after link integration

When link integration is applied, the values found after multiplication or joining can be used in the same way as in evaluative network analysis (Van Cuilenburg et al., 1988); see Figure 1.

After multiplication, the relation between concepts A and D consists of two parts, that via B and that via C. The strength of the part via B is $.5 \times .5 = .25$; the strength of the part via C is $.4 \times .4 = .16$. By joining these two parts, the average strength is computed, this is $(.25 + .16)/2 = .21$. The strength of the relation is especially relevant in the situation of causality. Here it gives an easily interpretable indication of the probability that A will result in D.

Exclusive knowledge

The results of empirical research are often restricted by certain conditions. These are conditions posed by the context in which the research has taken place. In research on women, the results are true for women in the population investigated. The results might also hold for men from that population, but this has not been investigated; therefore the results are exclusive for women. This fact has consequences when graphs based on different studies are integrated. The original graphs are based on studies that hold for a certain population. The graphs to be integrated should at least deal with the same population or show results that can be generalized to the same population. Sometimes such a generalization is not possible: what is found for women can usually not be generalized to men as well. In this case the integrated graph holds only for women. The graph is exclusive for women.

Such a condition (only women) must be taken into account. This can be done in two ways. One is to present two networks, one for men and one for women; this solution results in at least two graphs.[4] The other way demands that a formal characteristic be added to the knowledge graph which indicates when a relation can be used in representing the available knowledge and when it cannot. This is the inclusion or exclusion characteristic.

A knowledge graph contains in part relations that characterize or exemplify something. Such relations usually are not restricted to a certain group; they can be used in all situations. An extreme example is that the control exercised in an organization (PAR relation) is not related to whether the investigation was performed among men or women. Other types of relations, like the causal and the association relation, might be problematic; often there will be no arguments known for stating why a relation that is found to hold for one group also holds for another group.[5]

The inclusion or exclusion characteristic is formulated as an if-statement:

If (condition) then (relation).

When a relation is found to hold in a study of women, the statement would start: If (gender = woman) then (relation). Assuming a computer program is used for the integration of knowledge graphs, the investigator would have to indicate only whether the condition holds for the graph under construction or not; the program will take care of including the relations that are allowed and excluding those that do not satisfy the condition.

In the case of very small conditional networks which are part of a greater network, the investigator might include both situations in the network. See, in the example at the end of this article, the situations of high and low costs of measurement of the quality of labour performance.

Complex concepts

The concepts used as points in the graphs are based on the definitions employed by the original investigators, which can differ for several reasons. One is object relativity. This occurs when a concept can refer to different characteristics or is considered from different perspectives. Examples are the king in chess and the king of a country, or the bicycle as a means of transport for many people and as a “danger on the road” for car drivers. Such concepts should be recognized in the process of concept identification. A second difference arises when investigators define concepts that indicate the same characteristic in different ways. Here one investigator includes in a concept parts which another investigator does not. One cannot redefine such concepts unless the original investigator does. Therefore it is good practice to consult this investigator.

Another problem is that there are few relations between concepts, especially in cases where these are broad concepts. Such complex concepts can be unravelled into more basic concepts that are interrelated. This unravelling might in some cases also resolve the differences between concepts just indicated.

It is possible that concepts are linked to a simple concept within a complex concept. If such is the case, one might replace the complex concept by the simple ones and their relations. This is done by unravelling, which will allow a clearer representation of the knowledge graph, as more relations are possible. Below is an example for the complex concept “career perspective” (Figure 2).

It often occurs that only the parts of a concept that the investigator started with are linked to other concepts. Now this first concept needs to be unravelled. This unravelling of concepts is called “expansion of concepts”. To a certain degree, it is the reverse of concept integration.

The concept “career perspective” is represented in Figure 2; this concept is linked to the concepts “employee” and “job”, in both cases with a PAR relation. The concept is a complex one; in fact it consists of two parts.

Figure 3 shows how it can be unravelled into these two parts, viz. the concepts “career” and “perspective”, which are linked by the PAR relation.

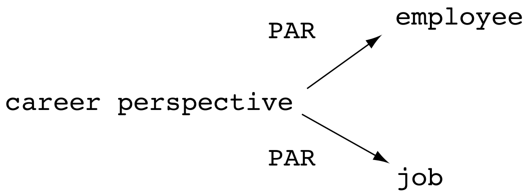

FIGURE 2
The concept “career perspective”

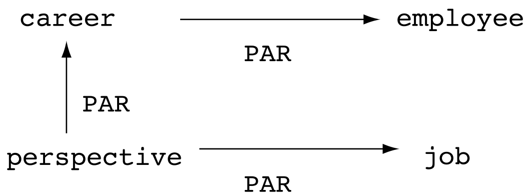

FIGURE 3
Unravelled concept "career perspective"

Here the concept "career" is linked to the concept "employee" (referring to the person) and the concept "perspective" is connected to the concept "job" that can be performed by any person. In this way the relation to another concept is made more specific, it is through the employee part or the job part of the unravelled concept.

The problem to be investigated is when this expansion of concepts is necessary, and whether guidelines can be formulated for performing this process. Starting-points might be found in qualitative research, where the construction of concepts is addressed.

Elaboration of a knowledge graph

The knowledge graph that received the most attention in Popping and Strijker (1997) is not the one based on the results found in six doctoral dissertations on labour markets, as there turned out to be hardly any concepts that were used in at least two of the studies. The final graph here consisted of many groups of concepts. Therefore a knowledge graph was presented based on theoretical ideas that had not yet been empirically tested. This graph was denser. It contained points, some of which were in fact frames, and the relations between these points. The frames referred to different theories. One of these is the control theory. For the moment, the other two theories (relation signal theory and function structure theory) are omitted. Now there are 16 points, and 14 relations between these points. One of the points is the frame, which consists in fact of 6 points and 5 lines. Taking the issues mentioned above into account, a knowledge graph results containing 14 points and

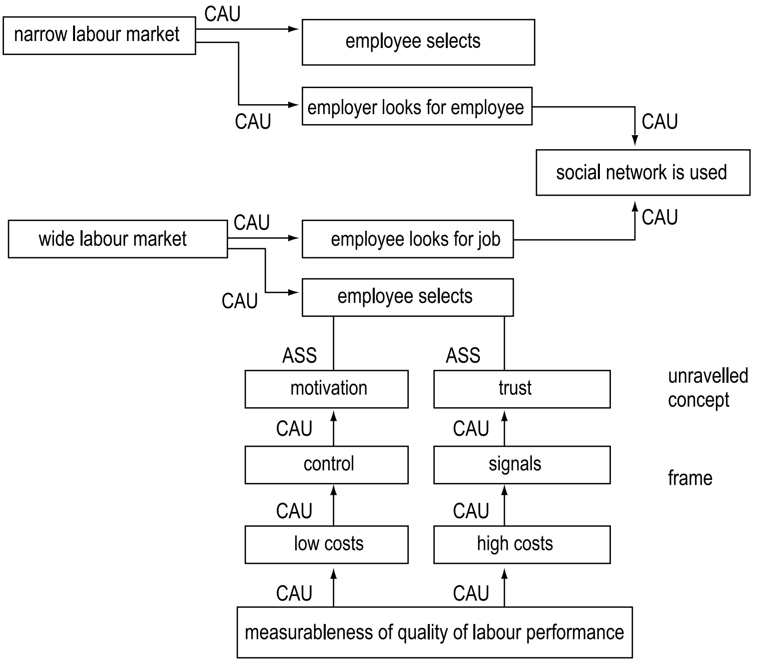

FIGURE 4
Knowledge graph of labour market theories according to control theory

14 relations (see Figure 4). The concept "control of labour relation" is unravelled, and the frame is linked to the graph through two different concepts.

Conclusions and discussion

The theory of knowledge graphs has been extended with considerations about implicit knowledge, strength of relations, exclusive knowledge and the unravelling of broad concepts. This should result in a more realistic representation of the empirically tested knowledge in a given domain. Taking these considerations into account, the knowledge graph on labour-market theory that was the basis of Popping and Strijker (1997) becomes more coherent.

The issues discussed are related to knowledge graphs. Before constructing such a graph, it must be decided on which (parts of a) text the graph should be based. For the coding, I followed the method that is also used in the evaluative network approach.[6]

Roel Popping teaches at the Department of Sociology at the University of Groningen, The Netherlands. He is currently working in the fields of sociology of knowledge, text analysis and knowledge representation by using graphs. Recent publications not mentioned in the references include: Popping, R. and Roberts, C. W. (1997) "Network Approaches in Text Analysis", in R. Klar and O. Opitz (eds) *Classification and Knowledge Organization*, pp. 381–9. Heidelberg: Springer; Popping, R. (2001) "Modal Auxiliaries in Text Analysis", *Bulletin de Méthodologie Sociologique* 74: 5–22. *Author's address:* Dr Roel Popping, Department of Sociology, University of Groningen, NL9712TG Groningen 31, Netherlands. [*email:* R.Popping@PPSW.RUG.NL]

Notes

1. The text-analysis process itself is not described here. In this process, however, one should stay with the concepts and relations used by the original investigator. This implies that, when a computer program is used for this coding, it must allow the user to follow the representational view on coding (Popping, 2000: 26). Here coding is performed according to the intended view of the original investigator. This is the opposite of the instrumental view, where in fact the process is automated and performed according to the view of the investigator constructing the graph.
2. Rahmstorf (1983) has defined 39 relations to handle nominal phrases. They allow exact representations, but due to their number they make the knowledge graphs unworkable.
3. Knowledge graphs also should not be applied to natural or common discourse. This will keep the investigator away from (ambiguous) sentences like "The man hit the dog with the stick". See Schank (1973) or Hoede and Willems (1989) for more details on this kind of approach.
4. Malrieu (1994) has proposed distinguishing such networks by using different colours. In this way they can be represented in one figure.
5. If, however, the relation also holds in the other group, the investigator will have to multiply the value of each relation by some constant, or replace it by some value which is correct for both groups. The similarity relation can be part of both situations distinguished here. This depends on the context in which such a relation occurs.
6. A computer program for constructing and manipulating knowledge graphs is available from the author upon request.

References

Bakx, G. J. T. A., Edens, R. G. and Hoede, C. (1987) "Knowledge Representation by Means of Knowledge Graphs: Application to Theories of Mathematical Modelling", I, Faculty of Applied Mathematics, Twente University of Technology (Memorandum 636).

Carley, K. M. (1986) "An Approach for Relating Social Structure to Cognitive Structure", *Journal of Mathematical Sociology* 12: 137–89.

Carley, K. M. (1988) "Formalizing the Experts' Knowledge", *Sociological Methods & Research* 17: 165–232.

Carley, K. M. (1993) "Coding Choices for Textual Analysis: A Comparison of Content Analysis and Map Analysis", in P.V. Marsden (ed.) *Sociological Methodology 1993*, pp. 75–126. Cambridge, MA: Basil Blackwell.

Carley, K. M. and Palmquist, M. (1992) "Extracting, Representing, and Analyzing Mental Models", *Social Forces* 70: 601–36.

Carré, B. (1979) *Graphs and Networks*. Oxford: Clarendon Press.

Hoede, C. and Weening, H. M. (1999) "Graph Theoretical Analysis as Tool for Validation and Conceptualization", in C. Van Dijkum, D. De Tombe and E. Van Kuijk (eds) *Validation of Simulation Models*, pp. 70–87. Amsterdam: SISWO.

Hoede, C. and Willems, M. (1989) "Knowledge Graphs and Natural Language", Faculty of Applied Mathematics, University of Twente (Memorandum 811).

James, P. (1992) "Knowledge Graphs", in R. P. Van de Riet and R. A. Meersman (eds) *Linguistic Instruments in Knowledge Engineering*, pp. 97–117. Amsterdam: Elsevier.

Malrieu, J. P. (1994) "Coloured Semantic Networks for Content Analysis", *Quality & Quantity* 28(1): 55–81.

Popping, R. (1998) "Knowledge Graphs in Group Learning", in Y. Waern (ed.) *Co-operative Process Management*, pp. 211–25. London: Taylor and Francis.

Popping, R. (2000) *Computer-Assisted Text Analysis*. London: Sage.

Popping, R. and Strijker, I. (1997) "Representation and Integration of Sociological Knowledge by Using Knowledge Graphs", *Social Science Information* 36(4): 731–47.

Rahmstorf, G. (1983) "Die semantische Relationen in nominalen Ausdrücken des Deutschen" [The Semantic Relations in Nominal Expressions in the German Language], unpublished doctoral dissertation, Johannes Gutenberg University, Mainz.

Rich, E. (1983) *Artificial Intelligence*. New York: McGraw-Hill.

Schank, R. (1973) "Identification of Conceptualizations Underlying Natural Language", in R. Schank and K. Colby (eds) *Computer Models of Thought and Language*, pp. 187–247. San Francisco, CA: Freeman.

Sowa, J. F. (1987) "Semantic Networks", in C. S. Shapiro (ed.) *Encyclopedia of Artificial Intelligence*, Part II, pp. 1011–24. New York: Wiley.

Stokman, F. N. and De Vries, P. H. (1988) "Structuring Knowledge in a Graph", in G. C. Van der Veer and G. Mulder (eds) *Human–Computer Interaction: Psychonomic Aspects*, pp. 186–206. Berlin: Springer.

Van Cuilenburg, J. J., Kleinnijenhuis, J. and De Ridder, J. A. (1986) "A Theory of Evaluative Discourse: Towards a Graph Theory of Journalistic Texts", *European Journal of Communication* 1: 65–96.

Van Cuilenburg, J. J., Kleinnijenhuis, J. and De Ridder, J. A. (1988) "Artificial Intelligence and Content Analysis", *Quality & Quantity* 22: 65–97.

Van den Berg, H. and Hoede, C. (1991) "A Knowledge Graph of Knowledge Representation Theory", Faculty of Applied Mathematics, University of Twente (Memorandum 980).

Yuen, H. K. and Richards, T. J. (1994) "Knowledge Representation for Grounded Theory Construction in Qualitative Data Analysis", *Journal of Mathematical Sociology* 19: 279–98.