
Faculty of Mathematical Sciences

University of Twente

University for Technical and Social Sciences

P.O. Box 217

7500 AE Enschede

The Netherlands

Phone: +31-53-4893400

Fax: +31-53-4893114

Email: memo@math.utwente.nl

MEMORANDUM No. 1500

Evaluating attitudes from texts

C. HOEDE

SEPTEMBER 1999

ISSN 0169-2690

Evaluating Attitudes from Texts

Cornelis Hoede

Faculty of Mathematical Sciences
University of Twente
P.O. Box 217
7500 AE Enschede, The Netherlands

Abstract

In this paper an example is given how the theory of knowledge graphs and the theory of social atoms can be used to evaluate the attitudes that actors have with respect to each other as far is evident from a text.

Key words: Knowledge graphs, social atoms, attitudes, texts.

AMS Subject Classifications: 05C99, 68D99, 92A20.

1 Introduction

The direct reason for writing this paper was a discussion within one of the sections of SISWO, the Dutch organization of social research, about a presentation by H. van den Berg and K. van der Veer on the evaluation of CETA. CETA is a text analysis system, where the outcome is a set of statements about the attitudes of actors towards each other, that result from the interpretation of a text by coders. The reader is referred to [1] for details of the CETA system.

The first weak spot of the system is the fact that live coders are used with their natural distribution of ability to interpret the text and the consequent distribution of interpretations. Even when all coders interpret the text correctly, there is variation in the actual interpretation, in particular when attitudes between occurring actors and issues as well as between one actor and other actors are to be estimated on a scale. The second weak spot again has to do with the live coders, who may mix up things in a way occurring rather often in policy studies. There the subject of study may be some system that should be described without any valuation whatsoever, so its actual functioning should be described and the policy discussion should be based on the input-output behaviour of the system, using as a criterium some multidimensional valuation function on the resulting outputs to determine the best set of inputs.

An example is the discussion about the enlargement of Schiphol airport. Before entering the discussion of the choice of the criterium, it should be clear which knobs can be turned to change the inputs and how they influence the outputs. If noise is an important output, it is not uninteresting to know that one input is the number of flights but that another input is the quality of the engines, that might be made less noisy.

The analogue of the mixing up of the actual system and the discussion about the control of the system, is the mixing up of issues that play a role in a text and the valuation of these issues by different actors occurring in the text as well as by the coders themselves. When coders have to determine the issues and have to estimate the valuations as well, they are easily liable to errors. In the following sections their task will be made easier by showing that a text can be prepared by knowledge graph theory and that the discussion about attitude can be made more transparent by the theory of social atoms.

2 Knowledge graph theory in a nutshell

A directed graph $G = (V, A)$ consists of a set of elements V , called *vertices*, and a set A of ordered pairs of V , called *arcs*. A is a subset of the Cartesian product $V \times V$. A *knowledge graph* is a directed graph of which the vertex set V represents concepts and the arc set A represents relationships between these concepts, that are of a certain *type*. The types available are restricted to only a few, that are considered to be basic in the sense that all relationships between two concepts can be reduced to graphs, containing these concepts, in which only basic types occur. The types form an *ontology*.

There are eight binary relationships and four n -ary or frame relationships. We will use the types EQU (equal), SUB (as in subset), ALI (alike, having things in common), CAU (causal), PAR (for attributes) and FPAR (for frame part).

For the description of systems the CAU-type of relationship is the important one. Verbs are represented by a vertex for the verb and CAU-links from subject to verb and from verb to object. This for transitive verbs. For intransitive verbs only one CAU-link is used from subject to verb. So MAN HITS DOG is represented by

$$\text{MAN} \xrightarrow{\text{CAU}} \text{HIT} \xrightarrow{\text{CAU}} \text{DOG}$$

and MAN SLEEPS by

$$\text{MAN} \xrightarrow{\text{CAU}} \text{SLEEP}.$$

For the description of attitudes the PAR-link is the important one. A careful distinction should be made between SUB-, PAR- and FPAR-link. The standard examples are

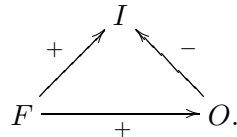
$$\text{HAND} \xrightarrow{\text{SUB}} \text{ARM}, \text{NICE} \xrightarrow{\text{PAR}} \text{GIRL}, \text{ANIMAL} \xrightarrow{\text{FPAR}} \text{CAT}$$

described by HAND is part *of* ARM, NICE is attribute *of* GIRL and ANIMAL is property *of* CAT. A lot of the mereological problems, making distinctions between various "part of" relationships, are due to the fact that the same word *of* is used in three essentially different situations. In the first example the molecules of the HAND are a *subset* of the molecules of the ARM. In the second example the GIRL is NICE because somebody thinks so, not because she has that as an objective property. In the last example ANIMAL belongs to the frame CAT, it is a property occurring in the definition of CAT. The FPAR-link can be seen as the inverse of the well-known ISA-link; a CAT ISA ANIMAL. For more

information on knowledge graphs, see the paper of James [4], although in that paper fewer types are considered.

3 Social atoms as generalisation of Heider's theory

Heider [3] has introduced *cognitive dissonance* in the following way. Given a focal person F , an other person O and some issue I , we consider the following simple labeled directed graph

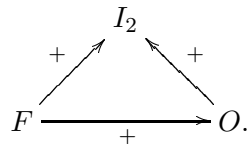


F likes O and has a positive attitude towards I . O , however, has a negative attitude towards I . This constitutes, according to Heider, a cognitive dissonance. He considers all eight combinations of (+)- and (-)-signs for this configuration and calls them consonant in case there is an even number of (-)-signs and dissonant in case there is an odd number of (-)-signs. This was taken as the basis of a theory of *balanced signed graphs* by Harary [2].

It was pointed out by Hoede [6], that the configuration is asymmetric. The dissonance occurs for F , not necessarily for O , who perhaps dislikes F . The (-)-sign from O to I is "as perceived by" F , who really is the *focal* person.

Another important remark is that there are three ways to cope with the dissonance. F may now start disliking O , bringing back consonance, or may change his positive attitude towards I or may even think his perception of O 's attitude towards I is wrong (how could his good friend have a different opinion?!). This remark is important for our discussion as it sheds some light on a possible *calculus of attitudes*, as is part of CETA. Dissonance resolution may take place in different ways and it may even be so that dissonance, as considered by Heider, is the most desired state for certain issues.

These are the elements of the theory of social atoms. Dissonances occur within the focal person and can be resolved in more ways than one. Moreover resolution of dissonance on one issue may cause dissonance on another issue. Suppose next to attitudes towards I_1 , with the configuration given before, we have the following configuration with respect to I_2 .



When resolution of dissonance on I_1 by F is done by changing his liking of O , then there will be dissonance on I_2 . As actual persons are interested in many issues, their actual behaviour in coping with dissonance may be very difficult to predict. This holds the more so, as people usually have different valuations for the same kind of dissonance. For our analysis these aspects of social atom theory should suffice.

4 The text and the actors occurring in it.

The example text was a New York Times editorial of 04-08-1981, page A14 and dealt with a rather well-known strike of air controllers.

HOLDING UP AMERICA

1. “May be we are crazy”, said Michael Fermon, a vice president of the striking Professional Air Traffic Controllers Organisation.
Then again, maybe the controllers, like most everyone else, would just like to work shorter hours for higher pay.
2. Whatever the merits of their case – and they appear to be dubious – the air controllers have no right to hold up the nation.
3. President Reagans tough threat to fire workers who are not back at work by Wednesday is appropriate.
A settlement that rewards them for illegally withholding vital services would be a serious mistake.
4. The 15.000 union controllers currently earn an average of \$34.000 a year. They had been negotiating with the Government for some months, asking for pay and fringe increases that would more than double their compensation.
In June, hours before a threatened strike, union leaders backed down. They accepted a package that would have raised their income by about 10 percent a year.
5. But the aroused rank-and-file members were not to be bought off so cheaply.
Air controllers bear responsibilities as great as those of airline pilots; so they feel they deserve salaries comparable to those won by the tough pilots’ union.
The controllers overwhelmingly rejected the settlement proposed by their union leaders and set the stage for yesterday’s walkout.
6. Although their work certainly requires discipline and creates stress, it is hard to feel much sympathy for the controllers.
There is no evidence that the work is debilitating.
At a time when other Federal employees are asked to accept a 4,8 percent raise, there is little justification for giving them more than twice that much.
7. But beyond that, the equities here are really beside the point.
The controllers have no legal right to promote their interests by damaging the national economy.
8. If President Reagan were now to sweeten the deal already cut in June, he would only be inviting other Government employees in key positions to exploit their leverage.

9. Living temporarily without regular air service is a heavy burden.

Restoring it on the controllers' terms could be a disaster.

As the main goal of CETA is to discuss attitudes, it seems best to locate the actors first. They are mainly persons but corporate actors are mentioned as well like PATCO, the Professional Air Traffic Controllers Organisation or the pilots' union. One person should not be forgotten. It is the journalist that wrote the text, who here and there clearly shows his attitudes.

We start by introducing letters for the actors

journalist	:	J
Michael Fernon	:	MF
AMERICA	:	A
vice president	:	V
PATCO	:	O
nation	:	A
President Reagan	:	R
government	:	G
leaders	:	L
members	:	M
pilots	:	P
pilots' union	:	U
employees	:	E .

Note that we identified AMERICA with nation, using our own background knowledge. We also consider MF and V to be equal, i.e. in a knowledge graph an EQU-link is chosen between MF and V . Other links that can already be indicated are

V	FPAR	O	Sentence 1.1
L	FPAR	O	Sentence 4.3
M	FPAR	O	Sentence 5.1
P	FPAR	U	Sentence 5.2
E	FPAR	G	Sentence 6.3 and Sentence 8.1.

Knowing that R FPAR G and G FPAR A is background knowledge not mentioned in the text. Also the equating of Federal with Government, in combination with employees, is using background knowledge.

Sofar we obtain the following graph,

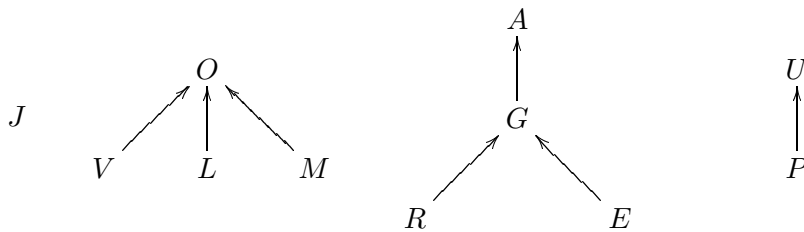


Figure 1

where all arcs are of type FPAR.

5 The attitudes coming forward in the text

After having located the actors, we can focus on their attitudes. The text is full of words referring to attitudes. Now note that we saw before that a social atom F has an attitude towards an issue I . There are three things involved, F , the attitude and I .

The first sentence “...WE...CRAZY” said MF ,... the three things are mentioned. MF is the focal person, CRAZY is the attitude towards WE, which is the issue here, namely O . We use some reasoning here about the word WE. $MF \equiv V$ must be part of WE so, as far as we know, $WE \equiv O$. It is the goal of knowledge graph theory to enable such reasoning automatically by defining *word graphs* for every word, but that is far beyond the scope of this paper, see e.g. Hoede and Li [5].

The structure is quite clear here:

$$MF \xrightarrow{\text{CAU}} \text{SAY} \xrightarrow{\text{CAU}} \text{“CRAZY PAR } O\text{”}.$$

Note the parenthesis that are used to express that a sentence, a frame, is said. In the Heider configuration $F \xrightarrow{+} I$ we might read $F \xrightarrow{\text{CAU}} \text{SAY} \xrightarrow{\text{CAU}} \text{“+ PAR } I\text{”}$ analogously. This rather complex way of describing attitudes will be shortened to $(F, +) \xrightarrow{\text{PAR}} I$, respectively $(MF, \text{CRAZY}) \xrightarrow{\text{PAR}} O$ in the example. Important is to remember who has the attitude.

We now find, i.e. the author as coder finds,

1. $(V, \text{CRAZY}) \rightarrow O$
 $(O, \text{LIKE}) \rightarrow \text{WORK SHORTER HOURS FOR HIGHER PAY}$
2. $(J, \text{DUBIOUS}) \rightarrow \text{MERITS}$
 $(J, \text{UNRIGHTFUL}) \rightarrow O \text{ HOLDS UP } A$
3. $(J, \text{TOUGH}) \rightarrow \text{THREAT}$
 $(J, \text{APPROPRIATE}) \rightarrow \text{THREAT}$
 $(J, \text{REWARDING}) \rightarrow \text{SETTLEMENT}$
 $(J, \text{ILLEGALLY}) \rightarrow \text{WITHHOLDING}$
 $(J, \text{VITAL}) \rightarrow \text{SERVICES}$
 $(J, \text{SERIOUS}) \rightarrow \text{MISTAKE}$
4. NO ATTITUDES FOUND
5. $(M, \text{CHEAPLY}) \rightarrow \text{BUY OFF}$
 $(O, \text{GREAT}) \rightarrow \text{RESPONSIBILITIES}$
 $(O, \text{DESERVED}) \rightarrow \text{SALARIES}$
 $(J, \text{TOUGH}) \rightarrow U$
 $(O, \text{REJECTABLE}) \rightarrow \text{SETTLEMENT}$

- | | | | |
|----|------------------------------------|---|------------------------------------|
| 6. | (<i>J</i> , DISCIPLINE REQUIRING) | → | WORK |
| | (<i>J</i> , STRESS CREATING) | → | WORK |
| | (<i>J</i> , UNSYMPATHETIC) | → | <i>O</i> |
| | (<i>J</i> , UNJUSTIFIABLE) | → | GIVING |
| 7. | (<i>J</i> , BESIDE THE POINT) | → | EQUITIES |
| | (<i>J</i> , ILLEGAL) | → | PROMOTING INTERESTS |
| | (<i>J</i> , DAMAGED) | → | DAMAGING |
| 8. | (<i>J</i> , SWEETENED) | → | DEAL |
| 9. | (<i>J</i> , HEAVY) | → | BURDEN |
| | (<i>J</i> , DISASTROUS) | → | RESTORING IT ON <i>O</i> 'S TERMS. |

All arcs are of the PAR-type.

Note that small changes to description by adjectives or adverbs are made. Also note that most attitudes are coded as coming from the journalist. For a coder to give a grading to the attitudes, on a five point scale say, seems rather difficult, but can be attempted once the above attitudes have been localized. The context need not play a disturbing role here. What is disturbing is that there is still great difference between the various adjectives and adverbs used. UNSYMPATHETIC is clearly expressing a negative attitude, but TOUGH may be seen as expressing a positive attitude or a negative attitude, depending on the desirability of being TOUGH. In SERIOUS MISTAKE the attitude is expressed more by the noun MISTAKE than by the adjective SERIOUS, which has a negative undertone but must be seen as describing a grading. CRAZY is a particularly funny example when coders are to give an interpretation.

The main methodological remark here is that our way of focusing on adjectives is expected to make evaluation of attitudes, if necessary, easier.

6 Linking attitudes with actors

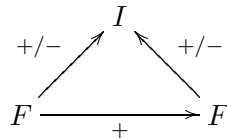
In Section 4 we discovered the actors and in Section 5 the attitudes expressed. We now have to link them in order to make an analysis of the cognitive dissonances possible.

Again we go directly at the core of the problem. The expressed attitudes concern certain concepts occurring on the right hand side of the PAR-relationships found in Section 5. We might first try to give a full translation of the text into a knowledge graph. But, due to the goal, we restrict ourselves to these concepts and use the text to relate them to the actors. The various analyses are

1. · O is explicitly mentioned
· “WORK SHORTER ETC.” is an actor independent statement.
2. · MERITS: “of their case” means “of the CASE of O ”
· O ’s action is explicitly mentioned
3. · THREAT: of R
· THREAT: of R
· SETTLEMENT: of R with O
· WITHHOLDING: of VITAL SERVICES by O
· SERVICES: to A
· MISTAKE: of R
- 4.
5. · BUY OFF: of O by R
· RESPONSIBILITIES: of O
· SALARIES: of O
· U explicitly mentioned
· SETTLEMENT: of R with O
6. · WORK: of O
· WORK: of O
· O explicitly mentioned
· GIVING: to O by R
7. · EQUITIES: of O
· PROMOTING INTERESTS: of O
· act by O
8. · DEAL: of R with O
9. · BURDEN: to A
· O explicitly mentioned.

Let us consider: $(J, \text{DUBIOUS}) \xrightarrow{\text{PAR}} \text{MERITS of the CASE of } O \cdot$

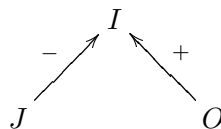
The journalist J thinks the merits of the case of O are dubious. Now, first, in all cases we assume that the attitude of the actor towards himself and his acts are positive, respectively consistent. So, in a Heider diagram,



In our notation and for our example we have

$(O, \text{POSITIVE}) \xrightarrow{\text{PAR}} \text{MERITS of the CASE of } O.$

Second, we clearly consider not O himself but the “MERITS of the CASE of O ”. So an attribute I of O is evaluated as dubious by J and, we assume, as positive by O . We have



and our problem is to say something about J 's attitude towards O . This is part of our intended "calculus of attitudes", on which later.

We will now use the Heider diagram representation with labels $+$, $-$ and 0 , the latter case if no positive or negative attitude is clear to the coder, in this case the author. I would, with these restrictions, choose

1. $\cdot V \xrightarrow{-} O \xleftarrow{+} O$, but the word MAYBE might be reason to skip this bit of information. Note that O , in the middle, is the issue here.
 - \cdot No attitude towards an actor involved.
2. $\cdot J \xrightarrow{-} \text{MERITS of the CASE} \xleftarrow{+} O$
 - $\cdot J \xrightarrow{-} \text{UPHOLDING A} \xleftarrow{+} O$.
3. $\cdot J \xrightarrow{0} \text{THREAT} \xleftarrow{+} R$.
 - $\cdot J \xrightarrow{+} \text{THREAT} \xleftarrow{+} R$.
 - $\cdot J \xrightarrow{+} \text{SETTLEMENT} \xleftarrow{+} R$.
 - $\cdot J \xrightarrow{-} \text{WITHHOLDING} \xleftarrow{+} O$.
 - $\cdot J \xrightarrow{+} \text{SERVICES} \xleftarrow{+} O$,
although withheld, the services of O are vital to A .
 - \cdot Due to the word WOULD no coding given.
4. \cdot No coding.
5. $\cdot M \xrightarrow{-} \text{BUYING OFF} \xleftarrow{+} R$.
 - \cdot Here the interpretation chosen is that O thinks O has great responsibilities, an example of self-esteem. No coding.
 - \cdot Again, for the same reason, no coding.
 - $\cdot J \xrightarrow{0} U \xleftarrow{+} U$.
 - $\cdot O \xrightarrow{-} \text{SETTLEMENT} \xleftarrow{+} R$.
6. $\cdot J \xrightarrow{+} \text{WORK} \xleftarrow{+} O$.
 - $\cdot J \xrightarrow{+} \text{WORK} \xleftarrow{+} O$, the journalist is positive about O 's work.
 - $\cdot J \xrightarrow{-} O \xleftarrow{+} O$.
 - \cdot Again we have a judgement for a hypothetical situation.
7. \cdot No coding, by the remark, literally, that "equities are beside the point".
 - $\cdot J \xrightarrow{-} \text{PROMOTING} \xleftarrow{+} O$.
 - $\cdot J \xrightarrow{-} \text{ACT} \xleftarrow{+} O$.
8. \cdot No coding, because of the word IF.
9. \cdot No coding as no explicit reference to O is made.
 - \cdot No coding, because of the word COULD.

Where actors were explicitly mentioned, we have direct attitudes. Most attitudes, however, concern issues of the other actor. We code them as follows.

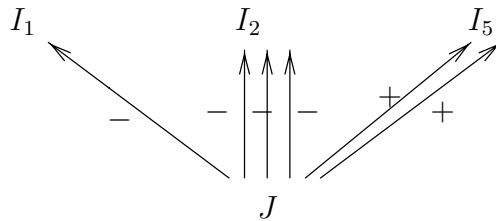
- I_1 : MERITS of the CASE, PROMOTING INTERESTS
- I_2 : UPHOLDING A, WITHHOLDING SERVICES, ACT
- I_3 : THREAT
- I_4 : SETTLEMENT, BUYING OFF
- I_5 : SERVICES, WORK .

Note that we identified some of the issues. We summarize:

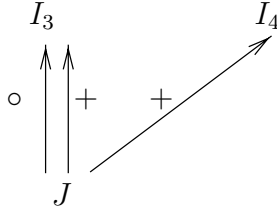
$$\begin{aligned}
 V &\xrightarrow{-} O \xleftarrow{+} O \\
 J &\xrightarrow{-} I_1 \xleftarrow{+} O \\
 J &\xrightarrow{-} I_2 \xleftarrow{+} O \\
 J &\xrightarrow{\circ} I_3 \xleftarrow{+} R \\
 J &\xrightarrow{+} I_3 \xleftarrow{+} R \\
 J &\xrightarrow{+} I_4 \xleftarrow{+} R \\
 J &\xrightarrow{-} I_2 \xleftarrow{+} O \\
 J &\xrightarrow{+} I_5 \xleftarrow{+} O \\
 M &\xrightarrow{-} I_4 \xleftarrow{+} R \\
 J &\xrightarrow{\circ} U \xleftarrow{+} U \\
 O &\xrightarrow{-} I_4 \xleftarrow{+} R \\
 J &\xrightarrow{+} I_5 \xleftarrow{+} O \\
 J &\xrightarrow{+} I_5 \xleftarrow{+} O \\
 J &\xrightarrow{-} O \xleftarrow{+} O \\
 J &\xrightarrow{-} I_2 \xleftarrow{+} O \\
 J &\xrightarrow{-} I_2 \xleftarrow{+} O \quad .
 \end{aligned}$$

We first focus on issue sets.

For J with respect to O we found



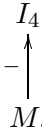
For I with respect to R we found



For O with respect to R we found



For M with respect to R we found



7 A calculus of attitudes

On all of these arcs we might have chosen a number from the interval $[-1, +1]$, indicating the strength of the attitude.

We face the problem of combining attitudes on issues. Let us just take the averages here of $+1$'s, -1 's and 0 's. We then find

$$\begin{aligned}
 J &\xrightarrow{-\frac{2}{6}} O \\
 J &\xrightarrow{+\frac{2}{3}} R \\
 (J &\xrightarrow{0} U) \\
 O &\xrightarrow{-1} R \\
 M &\xrightarrow{-1} R .
 \end{aligned}$$

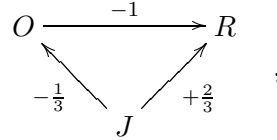
As M FPAR O , see Figure 1, we have found only three attitudes by considering mutual stands on issues.

We also had

$$\begin{aligned}
 V &\xrightarrow{-} O \\
 J &\xrightarrow{0} U \\
 J &\xrightarrow{-} O ,
 \end{aligned}$$

as direct attitudes. The journalist must be considered neutral towards U . The doubt cast by V on O can be left out of consideration as V FPAR O . Finally

$J \xrightarrow{-} O$ comes forward in the more detailed judgement based on the various issues. So ultimately we only find:



PATCO does not like President Reagan and the journalist favors President Reagan's stand and he is against the handlings of PATCO.

As far as the author is concerned, this is the result of evaluating the text. If more coders are present their results might be averaged. In CETA, next to the combination of attitudes towards issues of an other actor there is a calculus of the type: "the friend of my friend is my friend" and "the enemy of my friend is my enemy". In evaluating texts, like the one we considered, such reasoning afterwards is debatable. The path

$$J \xrightarrow{-\frac{1}{3}} O \xrightarrow{-1} R$$

yields, on "multiplication" of attitudes,

$$J \xrightarrow{+\frac{1}{3}} R.$$

This shows some consistence with the result

$$J \xrightarrow{+\frac{2}{3}} R$$

found directly, but slight additions or deletions in the text may seriously disturb such consistency. The advice should be to refrain from such "multiplication" of attitudes.

References

- [1] Berg, H. van den and C.G. van der Veer, *Computerized decision support systems and text analysis: evaluating CETA*. Vrije University Amsterdam, Faculty of Social, Cultural Sciences, Department of Social Research Methodology, Submitted to Quality and Quantity, (1999).
- [2] Harary, F., On the notion of balance of a signed graph, *Michigan Mathematical Journal*, 2, (1954), 143–146.
- [3] Heider, F., Attitudes and cognitive organization, *Journal of Psychology*, 21 (1946), 107–112.
- [4] James, P. *Knowledge graphs*, in Linguistic Instruments in Knowledge Engineering (R.P. van de Riet and R.A. Meersman, eds.), (1992), 97–117.

- [5] Hoede, C. and X. Li, *Word Graphs: The first set* in *Conceptual Structures: Knowledge Representation as Interlingua*, Auxiliary Proceedings of the Fourth International Conference in Conceptual Structures, Bondi Beach, Sydney, Australia (P.W. Eklund, G. Ellis and G. Mann, eds.) ICCS '96, (1996), 81–93.
- [6] Hoede, C., *Social Atoms*, Memorandum nr. 846, Faculteit der Toegepaste Wiskunde, Universiteit Twente, (1989).