BILETA

16th BILETA Annual Conference

April 9th - 10th, 2001.
University of Edinburgh, Scotland.

CODAS as a Tool for Jurimetrical Research.

LIA COMBRINK-KUITERS, RICHARD V. DE MULDER AND KEES van NOORTWIJK (Erasmus University, Rotterdam)

Presented by KEES van NOORTWIJK

Abstract

The initial aim of the CODAS-system, which has been developed by our Centre for Computers and Law (CCL), was to select documents from legal databases using conceptual techniques. As a spin-off of this system, an application was developed to mark student's assignments containing open questions. This system has been used for educational purposes for some years now and the results are very satisfying both for teachers and students.

Used in this way, the CODAS-system consists of two parts: one for finding similarities between digital documents - the *Fraud Finder* - and one for computer-assisted grading: the *Text Grader*. The *Fraud Finder* compares all the documents in a given set with each other and calculates a similarity score. The *Text Grader* sorts the documents according to how much they resemble certain selected example documents.[1]

Although the CODAS-system is now used for educational purposes, some pilots have been carried out to explore whether it could also be applied as a tool for jurimetrical research. As the results of these pilots are very promising, we have been encouraged to continue in this direction.

Our current research focuses on the following:

* A set of Family Law cases has been selected from Lexis/Nexis concerning both residence orders and contact orders. Some of the cases in this set closely resemble each other and we even found

cases that in fact appear more than once in the database. Using the *Fraud Finder* we have tried to locate these duplicates.

* Analysing judicial verdicts in a systematic manner often involves an accurate establishment of facts and circumstances appearing in a court case. Until now this analysis of the texts of judicial verdicts has been very time-consuming. If this process could be facilitated using conceptual techniques this would mean a great improvement in the method. The usefulness of the *Text Grader* has been tested in this respect to facilitate the knowledge acquisition process.

In this paper the results of the use of these two CODAS-applications will be presented. It consists of the following parts:

1. Introduction

2. The jurimetrical approach

3. Collecting two sets of cases from Lexis/Nexis

4. Using the Fraud Finder: Screening and Cleaning both sets

5. Using the Text Grader

6. Conclusion

1. Introduction

In the past few years the CCL has been involved in jurimetrical research into judicial decision-making. The aim of the studies was to provide an insight into the nature of judicial decision-making, as well as to develop an instrument to make such empirical research possible.

All kinds of verdicts can be used as a source for this kind of jurimetrical research, but the most interesting legal domains are those in which the legislator only provides vague guidelines for judges. The decision-maker has to come to a verdict on the basis of the facts and circumstances (we call them factors) of a specific case. Our jurimetrical research focuses on the relationship between these factors and the judicial decision. The domain of Family Law is, amongst others, very suitable for this purpose.

In order to determine which factors play a role in a judicial decision-making process, the texts of a number of cases on the same legal subject have to be analysed to establish the presence of certain factors. After having performed such an analysis on Dutch Family Law cases, we took the step of retrieving a series of UK-law cases on the subjects of 'Contact Order' and 'Residence Order' from the Lexis/Nexis databank.[2] One of the reasons for examining UK law was to make our research more accessible to a broader public.

2. The jurimetrical approach

In this section, we will give a general description of the jurimetrical method used to analyse texts of court cases. This method consists of the following steps.

Firstly, what we call the "legal item" for the two sets of cases has to be determined. The legal item for the first set of verdicts used here concerned the outcome, from the point of view of the father, in disputes on the residence order: had he been successful in his attempt to obtain or maintain the residence of the children? The legal item for the second set of verdicts concerned the question whether the non-custodial parent had been granted access rights by the court. It had to be determined

how the judge had decided in each case: "pro" or "con" on the legal item.

The next step would then be to develop a list of case factors, which would contain the characteristic features of the cases. The presumption is that each listed factor might have been of importance in the formation of the judge's decision. Usually, a large number of factors can be derived by means of a thorough text analysis. This is a very time-consuming procedure and special domain expertise is needed to accomplish this task.

The result of this analysis is a matrix containing, for all cases, both the factors and the judicial decisions on both items. The matrix shows which factors are present or absent in each case and in which way the judge decided (pro or con). These matrices contain all the research material on which the statistical analysis is based, to determine the extent and the way in which the case factors influenced the decisions. In order to achieve this, a statistical relationship has to be established between the presence and the absence of a factor in the cases and the way in which these cases were decided. This relationship is expressed in a particular association measure: the correlation coefficient. If this coefficient exceeds a previously fixed limit then it can be called significant. This gives rise to the presumption that the presence of a certain factor influenced the decision-making. This "influence" could have a "positive" or a "negative" direction, depending on the choice of the legal item (seen from the point of view of the father or the mother) as well as the way in which the factor is described. Where a statistical relationship is found, it is important to search for an explanation of this relationship, as the possibility of strictly coincidental links, despite the statistical precautions, is always present.

On the basis of these statistical relationships the cases can be ranked according to the probability that they were decided positively, i.e. according to their `strength'. In the listing, the `strong' positive cases are at the top and the `strong' negative cases are at the bottom.

Disadvantages of the method

The main problem of this method is the time consuming procedure to collect the necessary data before the statistical analysis can even begin. The Boolean searching procedure to collect the relevant verdicts and the manual analysis of these verdicts in order to establish the (potentially) relevant factors, demand much time and effort [Combrink and Piepers 1993], [Combrink and Piepers 1994] and [Combrink and Piepers 1995].

Firstly, a list of potentially important factors must be drawn up. This requires experience in the relevant domain. The formulation of the factors needs special attention to obtain the required information.

Secondly, we must look closely for the appearance of these listed factors in the cases. In all the cases it has to be assessed whether or not a factor is present. In order to ascertain this presence or absence, all cases have to be scrutinised several times. For the accomplishment of this task, a certain amount of legal knowledge is required. For example, it is possible that a certain article from a statute is not mentioned explicitly in the text of a case. The researcher must then be able to recognise a reference made to such an article by the appearance of certain (combinations of) words within the text of a verdict. The further preparation of the material which is necessary to make it suitable for statistical analysis is also labour intensive. All data with respect to factors and decisions has to be transferred from the working sheets into the computer. If we could automate any or all of these steps, in whole or in part, this would greatly alleviate the researcher's task.

To achieve this aim we will focus on the use of conceptual techniques, which were originally designed for the retrieval of documents from large databases. In these conceptual techniques, Bayesian statistical models and similarity measurements are used in order to calculate for each document the probability that it is relevant to a certain concept. Both these approaches have already

been implemented in a system called CODAS that is used to mark open question examinations and student assignments. We have now tried to use this system as an aid both in the selection of suitable cases and in the factor coding procedure. The underlying idea is to find relevant cases from a database, not by giving key words, but by showing the system examples of documents which are highly relevant and examples of documents which are not relevant. The program will evaluate the word use in these documents. With the help of the Bayesian statistical algorithm the system will gradually 'learn' to understand which documents the user is looking for. The documents are then ranked in such a way that the most relevant document will appear at the top of the list. The strength of Bayesian statistics is the use of information to improve the quality of the prediction of relevance. It would seem probable that the word use in documents provides vital information in this respect [Wildemast and De Mulder 1992] and [De Mulder 1995].

3. Collecting the two sets of cases from Lexis/Nexis

Our purpose was to find a sufficient number of cases concerning two strictly defined juridical items. The first set to be collected should contain cases on conflicts between parents with respect to the residence of their child(ren) after divorce. One of the parents asks the court to give a residence order. The second set should contain cases in which one of the parents asked the court for a contact order after divorce. These two subjects are closely related, which gives interesting possibilities for research.

To find the appropriate cases in the databank from 01-01-1990 onward we had to define a suitable keyword. We used 'residence order' and 'access' OR 'contact order' for this purpose. We retrieved 895 and 527 cases within the Lexis/Nexis database containing one of these words. As these numbers were too high and the cases included in it were not easily comparable, we decreased the number of cases in both sets by adding the word 'divorce' and by restricting ourselves to decisions given by the Court of Appeal. We also removed the cases concerning abduction from the residence set. This searching routine eventually resulted in two sets containing 159 and 87 cases respectively. After a first manual checking of the contents, a number of cases, especially from the residence set, were eliminated as being too deviant. Although they contained the keywords several times they were on different legal subjects.[3]

We started the experiment described in this paper with two sets containing respectively 109 cases (residence order conflicts between parents after divorce) and 82 cases(contact order conflicts between parents after divorce).

4. Using the Fraud Finder: Screening and Cleaning both sets

The first step we took was to carry out a preliminary screening of the two sets of cases. We wanted to make sure that every case appeared only once in each set. Furthermore, we wanted to know whether cases appeared in both sets and if so, if they had been registered under identical names. This was important because both, closely related, sets were merged to carry out several experiments.

For this step, we used the CODAS 'Fraud Finder': a program that compares all documents in a given set and produces a list of the 500 document pairs that are most similar. The similarity measure that is used is based on the word use in the documents [Combrink *et al.* 1999]. We used this program both on the two separate sets of cases and on a joint set containing all cases. The results were as follows:

* The Contact Order set contained 6 'duplicates': cases appearing twice under different names. This multiple occurrence was probably caused by the publication of a case in more than one periodical and/or location. This explanation became even more plausible when we observed that the two

versions of a certain case sometimes showed small differences, such as 'catchwords' added to the original text. Although these pairs of documents were not totally identical, the Fraud Finder program found them with similarity scores ranging from 94 to 99%.

* The residence order set contained 4 'duplicates'. Again, publication in different sources was the most probable cause. In one case, small typographical differences indicated that (parts of) the texts had been entered independently of each other. Once more, we found very high similarity scores for these documents (96-99%).

* When the two sets were joined, it was found that 24 cases were present in both sets. One of these appeared with a different name in the two separate sets.

After this screening procedure we had 153 cases at our disposal: 77 strictly residence order cases, 52 strictly contact order cases and 24 cases occurring in both sets. As some of these duplicate documents were among the largest of the respective sets, their removal can be of significant importance to the validity of the second part of the research.

Table 1.: Division of the cases

	Residence order	Not-residence order	Total number
Contact order	24	52	76
Not contact order	77	-	77
	101	52	153

These remaining 'clean' sets formed appropriate base material for the `Text Grader analysis' described in the next section.

5. Using the Text Grader

The CODAS 'Text Grader' is an application that can be used to rank a series of documents based on how similar they are to a series of example documents (and how dissimilar they are to certain counter-example documents). The program first produces an initial ranking. This ranking is based on *common features*: the document that has the highest number of characteristics that can be found in other documents as well is ranked highest. This initial score was originally intended to assist the user in finding a few example and counter example documents as quickly as possible. As soon as two or three of both these two types of documents have been found the so-called document scores can be calculated. These scores, based on the word use in the documents, then form the basis for a new ranking. For our purpose, the initial ranking results could also be used to check whether some cases had remained in our data sets which were highly dissimilar from all the others. As described in the next section this appeared not to be the case.

5.1 Ranking the cases

The use of the Text Grader started with an experiment to see if the initial score (as described above) can serve a different purpose as well, namely to indicate cases that are 'very uncommon', different from the rest of the set. For this, the initial scores were calculated for the cases in the two separate sets. These scores are always in the range 1..1000. For the contact order cases, the average initial score was 638.3 (standard deviation: 220.3), for the residence order cases this was 704.1 (standard deviation: 163.9). Initial scores lower than 2 times the standard deviation below the average score were considered extraordinary low. The cases with these scores were carefully inspected to see if

they had special characteristics. The results were as follows.

* Of the contact order cases, four had an 'extraordinary low' initial score. Apart from the fact that these four cases were all quite short (3 to 4 pages), no specific or uncommon characteristics were found. Two of the decisions were positive to the father, two were negative.

* Of the residence order cases, three had an 'extraordinary low' score. Here, the (short) length of the cases was not a common characteristic. One case was, on the opposite, very long indeed (52 pages), one was of average length (10 pages) and one was quite short (4 pages). With respect to the contents or the decision (positive or negative) again no common points for these three cases were found.

The conclusion from this first experiment was that the initial score did not point to specific and/or uncommon cases here. The only characteristic that probably accounted for the low values was the somewhat unusual size (quite small or, in one case, very large).

After we had looked at the initial ranking, we started to add codes to the cases to be able to perform the calculation algorithm. The set was divided in two groups: one containing the examples and the other containing counter examples. Due to the occurrence of cases in both sets, this was done in three different ways. Firstly, we separated the 101 residence order cases (including the cases occurring in both sets) from the 52 cases definitely not on residence orders. The same was done separating the 76 contact order cases (including the 24 cases occurring in both sets) from the 77 non-contact order cases. Finally we deleted the 24 cases concerning both subjects and only used the strictly residence order and strictly contact order cases, as each other's 'antagonists'.

5.2 Predicting the subject of a case

The aim of this conceptual experiment was to predict whether a case concerned a residence order or a contact order within the merged set. As explained above, the Text Grader uses the word occurrences within the documents to rank the cases. The more a case resembles the example documents, the higher it will appear in the ranking. Cases of which the word use strongly resembles the counter examples will appear low in the ranking.

5.2.1 The split-half method

We started with a rather traditional split-half approach. We divided the complete set of 153 cases into two parts: one part to build the model, the other part to predict the case with the help of this model.

Half of the cases were coded as example and counter example cases whereas the remaining half, being the cases to be predicted, were not coded. The example documents were marked with a +sign where it concerned a residence case and given a -sign where it was a non-residence case. The same was done with respect to the contact order/ non-contact order distinction.

It was expected that the cases concerning each subject that were not given a code would appear near the pro examples and the other cases (not concerning the subject) near the counter examples. We verified this hypothesis looking at a case's place in the ranking, taking the a-priori chance as a dividing line.

Residence order/ non-residence order

As said above, half of the 153 cases were used as examples. So 76 cases were coded according to their actual division in the whole set (101:52), therefore there were 50 examples and 26 counter examples. 77 cases remained to be predicted.

In order to be certain of a relevant result of this procedure, this result had to be better than the a-priori chance. This a-priori probability is 66 %, since we know that 101 of the 153 cases in the set are residence cases. It appeared that out of the 77 uncoded cases 55 appeared on the right side of the a-priori borderline, so 71 % of the cases (55/77) were predicted correctly using the split-half method. A random test, carried out as an extra check of the relevance of the result obtained, showed much worse results. Only 40 out of 77 cases (52 %) were ordered correctly.

Contact order/ non-contact order

Again, half of the 153 cases were used as examples. 76 cases were coded according to their actual division: 38 pro/38 con and 77 cases were left not coded.

The a-priori chance was 50,3 % (77/153). Out of the 77 not coded cases 53 appeared on the right side of the a-priori borderline. So 69 % of the cases were predicted correctly using the split-half method. Again the random test performed very badly. Only 33 out of 77 cases (43 %) were ordered correctly.

Residence order/ contact order

The split half technique was also tested after removing the 24 cases appearing in both sets.

There were 129 cases left: 77 exclusive residence order cases and 52 exclusive contact order cases.

Once again, half of the residence order cases were used as an example and coded + and half the contact order case were coded -. The 64 coded cases were used as examples and counter examples to predict the 65 non-coded cases.

The a-priori chance was 60 % (=77/129). Out of the 65 not coded cases only 37 were predicted correctly, which is 57 %. This is disappointing because we had expected an improvement on the results of the two previously described experiments. An explanation could be that from this set, although consisting of two more distinctive subjects, some very powerful cases had been deleted.

We also applied this split-half technique to a random set in this case. We only reached a score of 52 %, which was even lower than the a-priori chance of 60 %.

5.2.2: The (n-1)-method

Firstly, we added + and - codes to *all* the cases; + for the example case and - for the counter examples. With the help of the Text Grader a 'basic' ranking was calculated for each set. The aim of this (n-1) experiment was to predict the subject of a case - is it on residence orders or contact orders? - *while leaving out one of the cases each time*. This meant that the information about the subject of the case itself was left out for each case consecutively.

The Text Grader calculated the ranking for each case on the basis of the word-use in the remaining (n-1) cases. In this ranking, the positively coded cases appear on top and the negatively coded cases at the bottom. Each ranking was examined with respect to whether this particular uncoded case would then appear on the negative or on the positive side in the ranking. The a-priori chance was taken as a reference. So when an uncoded case appeared in the upper half above this line, between the positively coded cases, it was assumed to be a residence/contact order case. If it appeared in the lower half, it was considered to be a non - residence/contact case. This assumption was compared to the real situation. If these were matching, we considered that a case had been correctly predicted using the Text Grader.

Residence order/ non - residence order

The a-priori chance of this set is 66 %. The cases left out appeared 28 times out of 153 on the wrong side of the borderline. Therefore, using the Text Grader 82 % of the cases were ordered correctly.

Contact order/ non - contact order

The a-priori chance of this set is 50 %. Out of all uncoded cases 46 appeared on the wrong side of the borderline. 70 % of the cases were therefore ordered correctly.

Residence order/ contact order

The a-priori chance of this set of 129 cases is 60%. Without a code 24 cases were ranked on the wrong side of the line. So the result of this experiment was that 81% of the cases were ranked correctly.

5.2.3 The actual/reversed method

Again our starting point was a completely coded ranking. The aim of this experiment was to predict the subject of a case - whether it concerned a residence order or a contact order - by using its information alternatively both as a positive and as a negative case. In this approach, we calculated two rankings for both possible outcomes of each case. Both rankings were compared and the alternative that fitted best was selected. [4]After having made a choice between the two rankings, it was determined whether the selected alternatives corresponded to the actual or to the reversed code. If the actual code fitted best, then the case was defined as correctly predicted.

Residence order/ non- residence order

The a-priori chance of this set is 66 %. Of all the reversed cases 28 appeared to fit better. So 82 % of the cases were ordered better on the basis of their actual code.

Contact order/ non-contact order

The a-priori chance of this set is 50 %. Of all the reversed cases 49 appeared to fit better. So 68 % of the cases were ordered better on the basis of their actual code.

Residence order/ contact order

The a-priori chance of this set is 60 %. Of all the reversed cases 26 appeared to fit better in the alternative model. So 80 % of the cases were ordered better on the basis of their actual code.

Table 2. The overall results

set	n	A-priori (attainable gain)	Split half	random	n-1	Actual/ reversed
Residence/not residence	153	66% (+34%)	71% (+ 6%)	52% (-14%)	82% (+16%)	82% (+16%)
Contact/ not contact	153	50% (+50%)	69% (+19 %)	43% (- 7%)	70% (+20%)	68% (+18%)

Residence/contact	129	60% (+40%)	57% (- 3%)	52% (- 8%)	**80%** **(+20%)**	**81%** **(+21%)**

5.2.4 Some remarkable cases

The cases at the top and at the bottom of the original ranking were very long cases. For example the residence order case 96-12-13[5].

They were clear cases on each subject:

* Contact order 00-01-26[6]

* Contact order 96-03-05[7]

* Contact order 96-03-14[8]

* Residence order 92-05-22[9]

* Residence order 97-05-20[10]

Strong example case in the residence/ non-residence set were strong counter examples in the contact/ non-contact order set.

We also checked some cases in the middle of the ranking - the weak cases - such as

* Contact order 96-10-09[11] in which a grandfather was the applicant

* Residence order 95-07-04-03[12] in which the father did not wish to pursue his application for residence and handled about the grandparents wanting to see their grandchild

Some of the cases were remarkable in each of the rankings. For this reason we looked at these cases more closely to see if they were deviant and should, therefore, perhaps have been excluded from the set.

* *93-07-21 Occurring in both sets*[13]

This case concerns both subjects, so in both rankings it is considered as an example case. In the residence order ranking this is a remarkable case because it is predicted wrongly both in the n-1 and the actual-reversed ranking. In the contact order ranking, this case performs very well. It is stronger and was predicted correctly twice. This indicates that the case contains more contact order than residence order components. When the case is read this is confirmed. It is illustrated by the number of times the words are used: contact order 31 times and residence order 15 times.

* *96-02-01 Contact order case*[14]

In each of the rankings this is a weak case - especially in the contact order ranking - and it is always predicted wrongly. Reading the case makes it clear that the case is a contact order case, but the real conflict is about the man asking permission to stay and sleep in the wife's house during the weekend to see his child. That explains perhaps why it performs a little better in the residence ranking.

* *98-01-23 Residence order case.*[15]

In each of the rankings this is a weak middle-range case, especially in the residence ranking. This case is always predicted wrongly: twice in the residence order ranking and twice in the contact order ranking. After checking the text of the case it became clear that this case should not have been part of the collection at all, although it contains the words 'contact order'. The case does not really deal with a contact order or a residence order.

* *98-07-14 Contact order case*[16]

This case actually deals with an alimony conflict and performs badly in each of the rankings. Its ranking shows a remarkable change when the actual/reversed method is applied - from very high to very low - while it appears as a weak case in the n-1 ranking. This case should be removed from the set.

6. Conclusion

The Fraud Finder is very useful as an aid to detect duplicates and odd cases within a set. The Text Grader appears useful to divide sets of cases on two or more different subjects.

With the help of the Text Grader, the subject of a case is predicted better than would be expected according to the a-priori probability. Only the split-half method in the smallest set (residence/contact) performs worse than the a-priori probability benchmark. All methods do better on the real set than on a test set of randomly coded cases.

There is hardly any difference between the outcome of the (n-1) method and the actual/reversed method. In the residence order/contact order set the results are the best (shown boldface in table 2): half the attainable improvement was realised. The results in the second set - although higher in percentage: 20% and 18% respectively - are in fact worse, because a smaller part of the attainable gain was realised.

The incorrectly predicted cases are mainly the weak cases in the ranking. In the residence/non-residence set alone, the wrongly predicted cases are predominantly the ones just *below* the a-priori borderline.

Cases on top or at the bottom of the original ranking perform best. These cases could be used as example or counter examples to retrieve cases from a larger set. The dubious cases could be left out in future research.

Some cases attract attention because of the big change in their ranking after their coding is reversed. After a closer examination of the text, an explanation for this could often be found. Two of these cases did not belong in the set after all.

The Text Grader performs very well in separating the two different groups in sets in which cases of both types appear.

Finally, this experiment indicates that both the Fraud Finder and the Text Grader can be useful tools for jurimetrical research into judicial decision making.

References

Combrink-Kuiters, C.J.M. and P.A.W. Piepers 1993. `The Implementation of Predictive Capabilities into Legal Computer Systems'. *In 8th Bileta Conference Pre-Proceedings*. University of Warwick, Coventry, p. 63-72.

Combrink-Kuiters, C.J.M. and P.A.W. Piepers 1994. `Statistically Analysing Court Decisions on

Custody Disputes'. In *9th Bileta Conference Pre-Proceedings*. University of Warwick, Coventry, p. 65-72.

Combrink-Kuiters, C.J.M. and P.A.W. Piepers 1994. `Using Statistical Techniques as a Tool for the Analysis of Court Decisions. In *Proceedings of the 7th Australian Joint Conference on AI.* University of New England, Armidale, Australia, p. 602 and in *Poster Proceedings of the Conference*, p. 17-24.

Combrink-Kuiters, C.J.M. and P.A.W. Piepers 1994. `Knowledge Acquisition from Court decisions: a Jurimetrical Approach'. In *Conference Proceedings van SSIT 94.* Amsterdam.

Combrink-Kuiters, C.J.M. and P.A.W. Piepers 1995. `The Use of Information Systems in Research for the Aquisition of Knowledge'.*In 10th Bileta Conference Pre-Proceedings*. University of Stratclyde, Glasgow, p. 23-31.

Combrink-Kuiters, C.J.M, R.V. De Mulder and C. van Noortwijk. 'Analysis of Case Law: Measuring Similarity as an Aid to Coding Factors in Cases'. In: *The Law in the Information Society; Legal Documentation Administrative Innovation The Lawyer's Training and Role*. Florence, 2-5 December 1998, 12 pp.

Combrink-Kuiters, C.J.M., R.V. De Mulder, H. Elffers and C. van Noortwijk. 'Comparing Student Assignments by Computer'. In: *Cyberspace 1999: Crime, Criminal Justice and the Internet*. 14th annual Bileta Conference, York, 29th & 30th March 1999. Published in electronic form (CD-rom), 10 pp.

Mulder, R.V. De 1995. 'Probabilistic approaches to legal concepts'. In *Verso un Sistema Esperto Giuridico Integrale*, Tomo I, Ciampi C.et.al.(eds.), Milaan 1995, p. 125-140.

Mulder, R.V. De and C.J.M. Combrink-Kuiters. 'Is a computer capable of interpreting case law?' In: *Journal of Information, Law & Technology (JILT)*, Issue 1/1996, 21 pp. CTI Law Technology Centre, Warwick 1996. ISSN 1361-4169.

Noortwijk, C. van and R.V. De Mulder. 'The Similarities of Text Documents'. In: *Journal of Information, Law & Technology (JILT)*, Issue 2/1997, 10 pp. CTI Law Technology Centre, Warwick 1997.

Wildemast, C.A.M. and R.V. De Mulder 1992. 'Some considerations for the design of conceptual legal information retrieval sysytems.' In *Legal knowledge based systems, Information Technology and Law,* Jurix '92, p. 81-92. Vermande, Lelystad.

---

[1] Some of our findings regarding these applications and the conceptual techniques used in them were already published in our 1999-BILETA-paper and several JILT-articles (1/1996,2/1997).

[2] We would like to thank mrs Wiebiena Heestermans of the University of Warwick for her help and co-operation in using this data bank to retrieve a sufficient amount of cases.

[3] For example the contact order cases: HEARING-DATES: 6 DECEMBER 1996 PANEL: BUTLER-SLOSS, THORPE LJJ which is a residence order case and 5 FEBRUARY 1993.[1993] Fam 314, [1993] 3 All ER 596, [1993] 3 WLR 369, [1993] 1 FLR 598, [1993] 1 FCR 932, [1993] Fam Law 407 concerning a paternity test.

[4] For a full outline see [De Mulder and Combrink 1996].

[5] [1999] QB 18, [1997] 1 All ER 614, [1997] 3 683, [1997] RTR 275, 35 BMLR 174, [1998] 1 FLR 304, [1997] Fam Law

[6] HEARING-DATES: 26 JANUARY 2000 PANEL: SIMON BROWN, WARD, POTTER, LJJ

[7] [1996] 4 All ER 28, [1996] 3 WLR 506, [1996] 2 FLR 65

[8] [1996] 4 All ER 1023, [1996] 2 FLR 214, 31 BMLR 107

[9] Re L (Minors) No CCFMI 91/1490/F

[10] HEARING-DATES: 20 MAY 1997 PANEL: SWINTON THOMAS, SCHIEMANN LJJ

[11] HEARING-DATES: 9 OCTOBER 1996 PANEL: BUTLER-SLOSS, WALLER LJJ

[12] [1996] 1 FLR 158, [1996] 3 FCR 30, [1996] Fam Law 76

[13] [1993] 2 FLR 802, [1994] Fam Law 8

[14] [1996] 2 FLR 273, [1996] 2 FLR 496, [1996] Fam Law 605

[15] [1998] 2 FLR 656, [1998] Fam Law 659, [1999] 1 FCR 318

[16] [1998] 2 FLR 1131, [1998] Fam Law 660, [1999] 3 FCR 36