Neural Network Method in Law KIS-L  
(Knowledge in Space-Law)  
Vladimir Vrecion, Pavel Vrecion

ABSTRACT: The legal expert system KIS-L, based on the original neural network method that we have been developing since 1990 in Prague, is described. Some general ideas about the development of information and expert systems in the area of law are added.

1. THE DEVELOPMENT OF ESs IN LAW

Very generally we can distinguish three basic types of possible ESs in the area of law:

1) ESs where the prevailing instrument of formalization, modelling, algorithmization of legal thinking and/or reasoning is (relatively deep, sophisticated) logical analyses, logical formalization.

2) ESs based on a space structures especially on neural network methods, where information or knowledge is considered as elements (points, knots) and their relations in space, where these elements, their positions and relations are permanently changing. These ESs can in principle more dynamic than ESs sub 1/but they are modelling intellectual and/or information operations «more roughly».

3) ESs for more special legally regulated procedures where law prescribes some quantifications, quantitative evaluations especially in tax law (see the works of L.T. McCarty), family law – determination of alimony, etc. They use mainly simple arithmetic, algebraic procedures at algorithmization there but mathematically nontrivial procedures can also be used (e.g. from the mathematical theory of graphs see [3]).

Ad 1/Now the time is coming to concentrate our endeavours on developing practical broadly usable ESs for some important parts of legal reasoning based on systematic logical analysis (see especially [1, 2, 4] and the works of H. Yoshino, L. Philipps, E. Faneli and others). Such ESs will
always create the core or skeleton of sophisticated computerization in the area of law.

Ad 2/Such ESs can serve as aids especially for the creation of better legal empirical experience, for better association of legal and empirical facts. In future they will be important component of complex legal ESs especially for «more global» intellectual operations.

Ad 3/ESs of this type can already effectively serve in present practice. They can also be incorporated as a part of complex and/or combined legal ESs in the future.

At the present time in the whole area of AI there is an endeavour to combine the strengths and advantages of rule-based ESs (as sub 1) and neural network ESs. Neural network ESs need a lot of samples to shape («learn, train») themselves and they work in a rather transparent but dynamic way. The rule-based ESs need a lot of expert work and the user can know a relatively unchangeable algorithm for the creation of answers.

Usually it is done through bridges which provide interfaces between the ES and neural network. Shells like NEXPERT (Neuron Data), GURU (Micro Data Base Systems) enable linking in external routines. There are e.g. C libraries like OWL (Hyper Logic) and NEUROSOFT (HNC), neural networks like NEURAL WORKS PROFESSIONAL II (Neural Ware), NEURO SMARTS (Cognition Technology), EXPLORE NET (HNC) available. Some of them can be converted to C source code and connected with an ES.

Now it is already possible to begin a more concrete constructive discussion about the general architecture of mixed and/or combined legal ES. We should like to take part in such discussion (in a separate article for this journal). The creation and experiences with ES of the type sub 2/, now described are needed for such a project (earlier we have worked in the areas sub 1/, sub 3/and at the creation of «standard» legal ISs).

2. THEORETICAL BASES OF THE ES KIS-L

Our information about certain areas of human activity generally represent a set of elementary pieces of information, knowledge and the relation among them: we can call it information space. In law we can call it legal information space, where elements of information mainly are: legal norms and their parts, actual relevant facts and elements of their description and relations among them.

It is of course relative what we consider as element of information and/or knowledge. It depends on the level, scope and aim of the intellectual
operation (thinking, consideration) in the framework of the given information space. The described representation of the information space of ES KIS-L respects this fact by corresponding dynamic structuring of information elements and/or information space.

The basic principle of ES KIS-L is to represent information and/or information space in the 3 D net. 3 D can be determined by the definition of knots, their connections – edges and by the determination of rules according to which the net is shaped.

The knots represent the objects which consist of header, memo and of the list of connection to other knots. The header is a text information term, specification which denotes the meaning of the knot. Memo is full-text placed on the hard disk, describing in more detail information in the knot e.g., legal norm, scientific explanation, etc.

The connections are oriented lines - edges defined by the start knot, end knot and a length. The length of distance between the knots i.e., the length of the respective edge is a function of probability (intensity) of relation between connected knots; connection length equals: 1/probability -1. For example if probability if A then B is 1 then the length of the connection is 0, if this probability is 0.9 then the length is 0.1111, etc.

It can be instructive to compare this information net and/or space with similar forms known in physics. In principle the connection behaves like springs and knots like joints. When the distance between joints connected by the spring is lower than the length of the spring there raises a repelling force, when the distance is longer than the length of the spring then an attraction force is present. Between knots which are not connected there is a small and constant repelling force similar to electrostatic force. This force leads space net tendency to inflate. There exist many stable shapes in this environment. The optimum shape is the shape in which the sum of deformation energies of constricted (stretched) springs is minimal.

In such organised information 3 D nets there act (as we call it) spread probability charges. Such a charge is spread from activated 1-knots through connection - edges. The connection functions like resistance similarly to that of the spread of an electrical charge in a wire connection. In the knots into which a charge comes from several connections, the charges are multiplied and increase.

Knots are activated by questions. For instance the question «In which city does John live?» activates the knots «John» «live» and «city». Through connections – edges this information is spread through the net. Some information charges will be received by the knots if they are in the information net), «London», «Firenze», «Walker», «address» etc. Let us suppose that
the knot «London» receives twice a strong information charge. Once because London is a city, once because John’s name is Walker and Walker lives in London. In the knot «London» a high charge will be concentrated, that’s why «London» will be considered to be the most probable and correct answer. A knowledge base organised in such a way has several interesting features:

1. The most probable position of the knot which receives the highest charges is in centroid. The centroid is calculated from the positions of activated knots.
2. A knot information charge can be calculated (with some error) from the relative position of the knots in space only. Closely related information is stored in a smaller space (subspace), probability charges rapidly decreases as a function of distance from the source of the given probability charge [see the graph 1].
3. Because of the guaranteed sharp decrease of the information charge and because of the concentration of the information charge around a centroid, most of the information stored in the net can be omitted at the search for the answer of a question. In this way an effective pruning of the search inference algorithm of the ES can be achieved.
4. In such a system we work with the direction of information relations (oriented connections - edges) therefore the inference process can be aimed and hypotheses quickly created.
5. Because distance between knots is a function of probability, the distribution of respective probability in a net is also a function of these distances. The calculation of probabilities can be based on the calculation of minimal distances. The searched main information is very probably the knot where the sum of minimal distances from a given source of the information charge is minimal. The distribution of information charges (and/or probability in information - knowledge space can be seen in the graphs 2 and 3.

3. INFECTION ALGORITHM

1. A user asks a question
2. The question activates the knots in the knowledge space which contains terms from which the question is constructed.
3. The centroid is calculated from the position of these knots. The weights of terms (and/or relations) which are specified by the user on constructing the question are also taken into account.
4. A specific number of knots are selected from the information space (by specifying of selected knots or radius of a ball - centroid sphere - subspace).

5. During pruning distances are calculated from the centroid.

6. The first evaluation of relevant searched information can be done. The output is the list of sorted knots according to the distances.

An additional more exact deduction on the base of selected knots can be done:

1. The minimal distances to every source of information charge are found (the source of the charges are the knots specified in the question). The sum of these minimal distances is counted for each knot in the specific sphere (ball). The output is again a sorted list according to the sum of distances.

2. The minimal distances are calculated. Along the direct lines the probabilities are calculated. The direct lines are defined as the connection (edges) for which the minimal difference between distances from the starting and end knot is equal to the length of the connection. The probabilities are calculated as the spreading of the information charge. These steps are repeated for every source of the charge. If in the knot the charge comes from several sources, then the resulting charge of the knot is calculated (similarly as calculation of probability of the spread of disturbance in a circuit at a serial or parallel connection) as conjunction ($p_1 \land p_2$) or by disjunctive form ($p_1 + p_2 - p_1 \cdot p_2$). The output is a list of knots sorted out according to probabilities. This procedure can serve as a calculation of direct hypotheses.

3. When the direct hypotheses are calculated they can be further corrected by taking into account some parallel brandies of deduction. Let us suppose that two direct lines are connected with a connection between two knots which are a part of different direct lines. The information charges from such knots are distributed in parallely connected direct branches. The resulting charges are calculated as mentioned above.

4. The software utilities of the system

The programme of KIS-L is written in Turbo Pascal for Windows. The logical procedures for the organising of data bases and in the inference
algorithm are formulated in Boolean logic. The standard knowledge base (subbase) can contain 30,000 knots (each max. 32 kb) plus corresponding full text on disk (max. 32 kb) for each knot.

Basic features of the program and its use:

1. The program is controlled by means of menus, dialogue boxes and/or icons (for illustration see fig. 1 below).
2. The questions are created from the list of terms specified in the knowledge base.
3. The terms in a question can be weighted.
4. An optimisation of the knowledge space is automatically done.
5. Inference deduction operations necessary for the creation of answers are based on finding out and/or formulation: distance from the centroid, sum of minimal distances, direct hypothesis, parallel hypothesis.
6. The connections between knots (terms) are oriented. They are presented during the creation of a knowledge base according to concreteness and/or generality of the used terms. They can be changed in the KIS-L environment with scrollers.
7. Knowledge bases can be written either in the KIS-L environment or in an independent text processor.
8. The knowledge base can be presented in a graph sphere form.
9. Browsing through the knowledge base.
10. The used edit procedures have functions: add, delete, connect, disconnect knots, merge knots, edit memo (the edit memo has the following basic functions: search, open file, save, cut, copy, paste).
11. The knowledge base can be compiled from a text file (ANSI, OEM format).
12. The standard Windows help is available.

In the present form KIS-L can be used as an alternative to some applications of relational databases, databases and the full text method. The behaviour of complex systems can be simulated by changes in the methods developed for the spread of information charges to the spread of a values calculated from regression equations and by changing the definitions of connected lengths corresponding to correlations.

The system was originally developed for the area of law. But as the majority of ESs it can also have other applications e.g., in the area of engineering. In this modification, functions for common engineering calculations will be implemented.
5. Legal Applications

KIS-L structures and stores legal knowledge - legal norms, factual legal cases, legally relevant data, legal doctrinal knowledge by means of the described 3D information space. At first we had to create special information spaces (subspaces) for individual branches of law. They can be connected in one common information space.

The relations among individual pieces of whole legal information have of course a different and often permanently changing character. These relations are mainly logical, causal, empirical, statistical and their intensity generally can be described by the probability of the relation if A then B and by the length of the respective edge (the system allows further elaboration and algorithmization of these relations). As real legal knowledge is permanently changing the corresponding information space must reflect these changes. Therefore it must accept new or changing relevant legal information and permanently restructure and recount described relations among individual legal data (knots). In order to have such legal information space and on it base an up to date ES it requires the permanent flow of new relevant information into data base-information space, in the other words, permanent learning improving of the ES.

In reality the legal information space normally functions as an integral one but we can distinguish (or if it is useful separately construct) in this legal information space:

- subspace of normative information i.e., normative (state law) subspace. In the anglo-saxon systems we can include generalised norms there which are contained in precedents and in legal customs.
- subspace of factual legally relevant data especially from the cases of application and/or implementation of law i.e., of legal norms. This subspace is structured according to the branches of law and according to the structure of cases of application of law. In the framework of this main structure it can be structured further according to the application of individual legal norms and even their parts. This structuration represents an integral part of the legal information space but it can also be stored separately and serves as a help for the creation of the right questions.
- subspace of legal doctrinal information especially from scientific literature, legal commentaries, interpretation of law and/or of legal norms. (It is also possible to include there, but separately, information from police literature especially detective stories. It could be useful for the
determination of the possible inquiry version at the investigation of criminal acts.)

In order to create more intelligent ESs it is necessary to add some general linguistic information e.g., especially trees of common terms depicting their relations from the point of view of their semantic generality. Also basic general geographical information formulating relations among geographical terms is useful, etc.

KIS-L is working with all sorts of legally relevant information and counts with all their above mentioned mutual relations. It presents as an answer to questions all the closest pieces of information from permanently changing legal information relating to a question i.e., weighted terms in a question. By an iterative procedure i.e., by the specifying or adjusting of questions on the base of answers it is possible to get really relevant and complete information - answers from collected knowledge in an ES.

It is possible to say that the system simulates the first stages of our real thinking on a certain legal problem. At first we and/or our brain want generally to map the problem i.e., we find out relevant terms, notions, their basic relations which can play some role «around» our problem. Only after we try to refine this first «rough» knowledge.

The described ES can effectively help to find this «rough» («associative») knowledge, to map the legal relevant problem completely from the collected information in the information base-space. This completeness is practically useful because practice clearly shows that man, especially when he is pressed for time, can omit even important relevant relations «around» his problem when it is most necessary to consider the complex area of law or even all the legal system.

Of course we could test and implement the ES only gradually i.e., at first in limited areas of law. The experience has shown that the larger area of law, and/or all the legally relevant knowledge we had in the stored information space the more useful and relevant answers we got.

The relations among legal data information, pieces of knowledge generally depicted in KIS-L in the form of a mathematical space structure can be continually elaborated further, especially the relation in the mentioned subspace containing normative information can be deeper algorithmized by means of complex procedures of logical inference [see 1, 2]. Such software could be combined with the software of the described ES.
Graph 1. Shows relation between probability (axis X) and distance from the source of information (axis X). Distance is measured from the source of information charge.

Graph 2. Boolean And shows relation between distance from the centroid (axis X) and probability (axis X). Distance is measured from the centroid calculated from nodes which are sources of information charge. Incoming charges are calculated according serial distribution of probability.
The KIS-L has been sufficiently practically tested in the academic field and in practice. The results have shown that in practice the system can be very useful for analyses of standard phenomena (problems, acts) made by some specialists, who can ask proper questions and can properly evaluate the answers in an aimed dialogue, e.g., in the investigation of criminal acts, where, in the ES, as broadly as possible criminalistic knowledge and practical experience is stored including factual data about known criminal acts, their discovery and investigation.

From a general point of view of legal informatics according to our theoretical estimation, the experience with this ESs confirms that such systems will a useful part of future combined, complex legal ESs. There they will serve as a part for the first, "rough", probabilistic orientation in a legally relevant problem. We are preparing a separate article about the architecture of combined legal ESs for this journal.

The software of KIS-L could be combined with some programs of general AI especially for linguistically more sophisticated dialogue with the computer i.e., in a more friendly language form.

**Graph 3.** *Boolean Or shows relation between distance from the centroid (axis X) and probability (axis Y). Distance is measured from the centroid calculated from nodes which are sources of information charge. Incoming charges are calculated according parallel distribution of probability.*
REFERENCES


