

Semantische Netze und ihr Einsatz in juristischen wissensbasierten Systemen

Alexandr Kurbatski, Vasily Cheushev, Julia Döring

Belarussische Staatliche Universität
Pr-t Nezavisimosti,4, 220030 Minsk, Weißrussland
kurb@unibel.by, cheushev@niks.by

Institut für Rechtsinformatik
Universität des Saarlandes, Postfach 151150, 66041 Saarbrücken
julia.doering@diedoerings.de

Schlagworte: Semantische Netze, Wissensdatenbank, wissensbasierte Systeme, semantische Darstellung von juristischem Wissen

Abstract: Es werden Grundlagen des semantischen Netzes, einige Fragen der Bildung der wissensbasierten Systeme auf Grund der existierenden wissenschaftlich-technischen Basis, sowie potenzielle Möglichkeiten deren Anwendungen bei dem Modellieren von juristischem Wissen betrachtet.

1. Die Vision eines semantischen Netzes (Semantic Network)

Zuerst stellen wir kurz Grundlagen des semantischen Netzes (SN) dar. Die Basis des SN und eine Wissensdatenbank auf ihrer Grundlage bilden Objekte, deren Zusammenhang und Regeln der Schlussfolgerung.

Als **Objekt** im SN kann man ein beliebiges Element der Informationsressourcen (IR) darstellen. **Objekte** des semantischen Netzes verbinden sich miteinander mittels verschiedenartiger **Beziehungen**, die alle notwendigen semantischen Verbindungen zwischen den Objekten des SN widerspiegeln. Beziehungen sind selbst Objekte des SN und können auch miteinander mittels anderen Beziehungen verbunden sein. Eine wichtige Kategorie der Objekte gründen **Fakten**, die bestimmte Gruppen der Objekte und deren Beziehungen zusammen verbinden.

Jedem Objekt wird ein **Typ** mit Hilfe einer der Beziehungen zugeschrieben. Jedes Objekt des gegebenen Typs übernimmt nach einer Hierarchie alle Eigenschaften (Merkmale), die für die Objekte seines Typs vorbestimmt sind. Die Übernahme der Merkmale des vorhandenen Objektes bezeichnet man als **Vererbung**.

Die wichtigste Rolle in einer Wissensdatenbank spielen **die Regeln der Schlussfolgerung**, die eine nicht offensichtliche Information durch die im SN existierenden offensichtlichen Beziehungen zwischen den Objekten offenbaren können. So kann man zum Beispiel einige Objekte durch ihren Zusammenhang mit anderen Objekten mittels Kaskadenfolgerung nach der Kette der Beziehungen finden. Dadurch wird implizites Wissen transparent und neues Wissen, welches aufgrund logischer Schlussfolgerungen gewonnen wurde, abgeleitet.¹

2. Das wissenschaftliche Paradigma

2.1 Die traditionelle Einstellung

Die meisten in der Praxis eingesetzten Datenbanksysteme sind für relationale Datenbanken konzipiert. Das kann man kurz auf folgende Weise charakterisieren. Für die Informationsressource (IR) des Systems wird ein relationales Datenbankmodell entworfen. Weiter wird für jede Art des Objektes der IR und jeder Aufgabe, die mit der Bildung der Komponenten der IR, ihr Begleitung und Darstellung verbunden sind, Verwaltungssoftware entworfen.

Mit anderen Worten werden die Strukturen der Komponenten der IR, deren Verknüpfungen und andere Metainformationen in die Software des Systems fest integriert.

Dieses Modell hat gewisse Vorteile:

Eine solche Methode ist einfach zu verstehen und ist gut erprobt. Letztendlich stehen dem Benutzer des Systems nur jene funktionellen Möglichkeiten zur Verfügung, die im System vorgesehen sind.

Im Grunde genommen, sind relationale Datenbanken derzeit immer noch die am meisten verbreitete und ausreichend wirksame Datenbankform für Anwendungen, bei denen den Daten und den Aufgaben nach ihrer Bearbeitung eine gewisse Steifheit natürlicher Art eigen ist. Zum Beispiel, wenn die Datenstrukturen und Such- und andere Funktionen des Systems für viele Jahre voraus reglementiert sind.

Andererseits, bereitet dies auch Nachteile und begrenzt potentielle Möglichkeiten des Informationssystems:

¹ *Cheushev*, Model "Diagram of solutions" for organizing data in information resources, The problems of legal informatization 1/2005, Minsk, Belarus.

Auf der Stufe der Datenmodellabbildung ist es unmöglich, alle Informationsanforderungen vorzusehen, die der Endbenutzer benötigen könnte, und selbst wenn es möglich wäre, hätte eine Realisierung allergrößte Entwicklungskosten und Ressourcen gefordert.

Als Folge sind das Informationsangebot und die Informationsnachfrage nur selten übereinstimmend: vielmehr gibt es einerseits häufig eine Nachfrage nach Informationen, die nicht angeboten werden, also ein Informationsdefizit, und andererseits gibt es häufig ein Angebot an Informationen, die nicht nachgefragt werden, also einen Informationsüberschuss.

Jede Änderung der Datenmodellabbildung erfordert eine Änderung der Software und als Folge steigen die gesamten Kosten des Projektes und die Entwicklung des Systems wird aufgehalten.

In der Praxis häufen sich die Fälle, in denen ein relationales Datenbankmodell eine unangemessene Lösung für Entwicklung eines Informationssystems ist. Als alternative Lösung kann man eine Veränderung der Projektsparadigma für die Versorgung der hohen Flexibilität des Informationssystems, sowohl in Beziehung zu der Organisation der Daten, als auch in Beziehung zu der Realisation der Funktionen vorschlagen. Dafür kann nach unserer Ansicht der Apparat der semantischen Netze von großem Nutzen sein.²

2.2 Das semantische Netz

In einer Wissensdatenbank des semantischen Netzes werden Objekte und deren Verknüpfungen auf der Basis der Grundprinzipien der Konstruktionen des SN geschaffen. Eine der Besonderheiten besteht darin, dass Metabeschreibung der Objekte und deren Verbindungen (Metadaten des Systems) nach denselben Prinzipien gebaut werden und einen untrennbaren Bestandteil des SN bilden.

Die Einheitlichkeit der Algorithmen für eine große Menge verschiedener Objekte lässt es zu, standardisierten Programmcode mit einer hohen Effektivität und einen breiten Satz typisierter Verwaltungs-, Such-, Navigations- und anderen Funktionen zu entwickeln.

Andererseits ist es möglich für jede Art der Objekte an den benötigten Stellen Programmcode zu definieren, um eine Realisation der spezifischen Funktionen für diese Objekte zu ermöglichen.

² *Kashinsky, Kurbatski, Cheushev, Doering*, About creation of legal knowledge bases, The problems of legal informatization 1/2006, Minsk, Belarus.

Bei solcher Einstellung zum Projektieren können neue Arten von Objekten und deren Verbindungen innerhalb weniger Minuten geschaffen oder undefiniert werden, dabei bekommt der neu geschaffene Typ des Objektes „fix und fertig“ die gesamten typisierten funktionellen Eigenschaften der Verwaltung der Objekte, der Such-, der Navigations- und anderer Funktionen.

Eine weitere Besonderheit besteht darin, dass hierarchische semantische Beziehungen der Komponenten der IR sowie Metadaten des Systems durch eine Datenstruktur „Diagramm“ repräsentiert werden können. Die Knoten des Diagramms stellen dabei die Objekte dar. Und Beziehungen zwischen den Objekten werden durch die Kanten des Diagramms realisiert. Bäume, die zu den meist verwendeten Datenstrukturen in der Informatik gehören, sind im Grunde genommen Spezialfälle von Diagrammen. Dabei erlaubt ein Diagramm im Vergleich mit einem Baum viel effektiver und kompakter Abbildungen von solchen hierarchischen Datenstrukturen wie zB Datenkatalogen und Datenverzeichnissen zu verwirklichen.³

Die Knoten des Diagramms sowie andere Objekte verbinden sich mittels Kanten verschiedener semantischer Beziehungen. Dadurch entsteht ein Semantisches Netz der Informationsobjekte.

3. Die Möglichkeiten

Diese Methode erlaubt viele Probleme, mit denen sich Experten und Wissensingenieure auseinander setzen müssen, schon auf dem Niveau der formalen Modelle (des logisches Entwurfs) aufzulösen. Zur Zahl der wichtigsten davon gehören zB Abbildung einer hierarchischen Struktur (mit der Zulassung, dass Objekte in viele Hierarchien gleichzeitig eintreten können) und hierarchische Vererbung der Eigenschaften (sowohl nach unten als auch nach oben), Einordnung der Gruppen der Objekte (für ein und das selbe Objekt sind verschiedene Einordnungen zulässig), das Problem der Synonymie (das Problem wird im weitem Sinne verstanden: es sind nicht nur dieselbe Definition und verschiedene Bezeichnung eines Begriffes, sondern auch dieselbe Benennung und verschiedene Bezeichnung eines Objektes sowohl in einer Sprache, als auch in mehrere Sprachen).

3 Siehe FN 1.

Auf solche Weise bietet die Entwicklung eines Informationssystems auf Grund der semantischen Netze mehr Flexibilität im Manipulieren von Daten und reichere funktionale Möglichkeiten bei der gleichzeitigen Senkung der Entwicklungskosten.⁴

4. Von juristischen Datenbanken zu juristischen wissensbasierten Systemen

Ein Übergang von juristischen Datenbanken zu juristischen wissensbasierten Systemen, ist allem Anschein nach dialektisch unvermeidlich, obwohl er in verschiedenen Formen verwirklicht werden kann.⁵

Im Bereich des Wissensmanagements spricht man heute von einem intelligenten Zugriff auf strukturiertes und vernetztes Wissen: Intellectual Information Retrieval statt Konstruktion von Expertensystemen.⁶

Ein grundsätzliches Problem bei dem Modellieren von juristischen Wissen und dessen semantischen Darstellung liegt in der empirischen Natur des Rechts. Dies bedeutet im Konkreten, dass viele juristische Begriffe keine eindeutige Definition (im mathematischen Sinne) besitzen und bekannte rechtliche Ontologien, die üblicherweise in Taxonomien, also in einer Baumstruktur organisiert sind, nur begrenzt (wieder)verwendbare sind.⁷ Es existieren weitere Ansätze, Wissensmodelle des Rechts aufzubauen und komplizierte hierarchische Strukturen des Rechts wiederzuspiegeln.⁸ Und *Semantisches Netz kann unserer Ansicht nach als einer davon betrachtet werden.*

Auf diese Art und Weise können zB semantische Darstellung von Rechtsnormen für die Analyse der rechtlichen Versorgung der problematischen Situationen sowie mehrsprachiger Zugang zur rechtlichen Informationen und andere im juristischen Alltag wichtige Aufgaben betrachtet werden.

4 Siehe FN 2.

5 Siehe FN 2.

6 *Staab, Studer, Schnurr, Sure*, Knowledge processes and ontologies. IEEE Intelligent Systems 01/02 2001.

7 *Doering, Herberger*, About legal ontologies, The problems of legal informatization 2/2006, Minsk, Belarus.

8 Siehe FN 6.

5. Die existierende wissenschaftlich-technische Basis

Zurzeit werden in der Belorussischen Staatlichen Universität Konzepte erarbeitet, die sich bereits heute zumindest ansatzweise technisch verwirklichen lassen. Es handelt sich um zwei Komplexe der Entwicklungswerke für die Konstruktion der wissensbasierten Systeme auf Grund der semantischen Netze, die sich voneinander durch die Mechanismen der Aufbewahrung der Daten, des Niveaus der Flexibilität und der Bedienungen der Anwendung unterscheiden.

Der erste Komplex erlaubt, eine Wissensbasis mit der maximalen Freiheitsstufe auf dem Niveau der „Wissensingenieure“ und „Experten“ zu schaffen. Man kann solche Systeme als das Werkzeug der Forschung und der Modellierung der semantischen Netze vor ihrer Konstruktion auf Grund des zweiten Komplexes benutzen.

Den zweiten Komplex bilden Algorithmen und Software der Entwicklung der wissensbasierten Systeme.

Die Idee besteht darin, dass eine Wissensbasis, außer Bildung „eigener“ Objekte und deren Verbindungen, auch auf ein relationales Datenbankmodell des Informationssystems eingestellt werden kann. Dadurch wird ein semantisches Netz mit dem Heranziehen schon existierenden Daten und deren Bearbeitung mit allen eigenen funktionalen Möglichkeiten der Wissensbasis entstehen. Mit Hilfe dieser Technologien kann man sowohl neue Informationssysteme entwickeln, als auch schon existierende modernisieren. Diese Methode fordert keine ernste Reorganisation der Datenstrukturen, was ermöglicht, bestehende Wissensbasen mit anderen Mitteln der Arbeit mit dem Informationssystem zu integrieren. Beide Komplexe verwenden SQL-Standard und spezielle Verwaltungssoftware für die Aufbewahrung der Daten und die Ausführung der Informationsanforderungen. Die ganze „Zwischenschicht“ der Datenbearbeitung stützt sich auf die XML-Technologie, die Benutzeroberfläche wird durch die Internet-Technologie verwirklicht.

Diese Technologien werden bereits in einem Monitoringsystem des Bildungs- und Kulturministeriums der Republik Belarus angewendet, um Wissen über Lehrmethoden aus heterogenen Datenbanken zu integrieren. Das System ist ein Echtzeitsystem und umfasst Informationen über mehr als 1400 Schulen, 1500 Lehrbücher und 1700 Lehrprogramme landesweit.

6. Fazit

Eine traditionelle Einstellung zur Projektierung auf Basis des relationalen Datenbankmodells erlaubt meistens nicht, ein flexibles und durchsichtiges Informationssystem zu verwirklichen.

Dafür kann nach unserer Ansicht der Apparat der semantischen Netze von großem Nutzen sein, was sich erfahrungsgemäß in der Belorussischen Staatlichen Universität bestätigt.

Aus diesem Grund betrachten wir den Apparat der semantischen Netze als einer der Möglichkeiten juristisches Wissen zu modellieren und einen intelligenten Zugriff auf strukturiertes und vernetztes Wissen zu verwirklichen.

Wahrscheinlich ist eine der Hauptfragen, inwieweit die heute existierende wissenschaftlich-technische Basis mit materiellen Ressourcen unterstützt werden kann, um viel stärkeren Praxisbezug im Bereich des juristischen Wissensmanagement zu gewährleisten?