

Proseminar Anthropomatik: Ausnutzung von Modellwissen beim probabilistischen Schlussfolgern

Benjamin Halbrock, E-Mail: kontakt@benjamin-halbrock.de

Zusammenfassung—Dieses Paper gibt eine Einführung in verschiedene Ansätze zur Abbildung und Nutzung von Hintergrundwissen zum probabilistischen Schlussfolgern. Betrachtet werden insbesondere die Object-Oriented Probabilistic Relational Language und Probabilistic-OWL auf Basis von Multi Entity Bayesian Networks.

I. EINLEITUNG

Die Integration von Technik in den Alltag nimmt immer weiter zu. Wir lassen uns vom Navigationsgerät durch den Straßenschwung führen, benutzen das Smartphone zur Informationsfindung und benutzen Fahrassistentenprodukte zur Unfallvermeidung. Immer mehr der Informationen auf die diese Geräte zurückgreifen werden automatisiert aus unsicheren teilweise fehlerhaften Daten gewonnen oder müssen direkt vom Gerät erkannt werden. Da Sensoren Messfehler liefern können und Menschen Fehler machen, kann den Informationsquellen nicht vollständig vertraut werden. Die Informationen müssen mit einem Unsicherheitsfaktor versehen werden. Ein weiteres Problem ist, dass sich die reale Welt auf Grund ihrer Komplexität nicht exakt im Computer abbilden lässt. Es gibt verschiedene Strategien die Sensordaten und vordefiniertes Modellwissen, mit Unsicherheiten, zu einem angepassten Weltmodell zu fusionieren. Auf Basis dieses Weltmodells lassen sich dann Situationen analysieren und Schlussfolgerungen und Entscheidungen treffen.

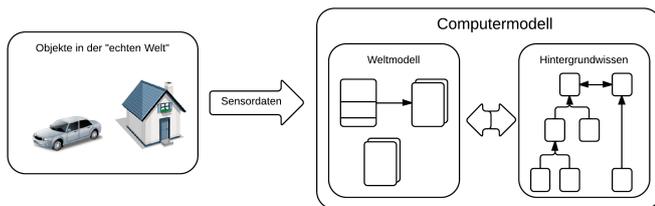


Abbildung 1. Reale Gegenstände werden über Sensoren erkannt. Diese Sensordaten werden mit Hilfe des Modell- / Hintergrundwissens als Objekte ins Weltmodell eingepflegt (nach: [1])

In Abbildung 1 erkennen Sensoren ein Auto und ein Haus. Über den Zusammenhang der erkannten Objekte können die Sensoren jedoch keine Aussage treffen. Dieser muss als Hintergrundwissen vorliegen. Ein Entfernungsmesser in einem Auto kann zwar feststellen, dass der Abstand zu einem Haus abnimmt, wichtig zur Einschätzung der Situation ist aber ob die Straße nur dicht am Haus vorbeiführt, gerade eingeparkt werden soll oder das Auto außer Kontrolle geraten ist. Es sollen Fragen wie: "Mit welcher Wahrscheinlichkeit wird gerade eingeparkt?" oder "Wie sicher ist das Überqueren der

vor mir liegenden Kreuzung im Moment?" beantwortet werden können.

Was eine Ampel ist, was ihre Farbe bedeutet und wann überhaupt abgebogen werden darf lässt sich in Ontologien modellieren. Aussagen von Ontologien sind immer absolut, sie kennen keine Unsicherheiten. Daher lässt sich mit ihnen keine Schlussfolgerung zur Frage "Wie sicher ist das Überqueren der vor mir liegenden Kreuzung im Moment?" treffen.

Mit einem Bayesschen Netz, welches zur Modellierung und Auswertung bedingter Wahrscheinlichkeiten genutzt wird, ließe sich die Frage "Wie sicher ist das Überqueren einer konkret definierten Kreuzung?" beantworten. Da ein Bayessches Netz keine allgemeinen Aussagen zu ähnlichen Situationen treffen kann, lässt sich zur konkret "vor mir liegenden Kreuzung" ohne Bilden eines neuen konkreten Netzes auch keine Schlussfolgerung ziehen.

Das Objektorientierte Weltmodell, Multi Entity Bayesian Networks, Probabilistic-OWL und die Object-Oriented Probabilistic Relational Language führen aus Ontologien und Bayesschen Netzen bekannte Konzepte zusammen, um Modellwissen zum probabilistischen Schlussfolgern nutzen zu können.

II. GRUNDLAGEN

A. Ontologien

Ontologien sind die formale Definition der Modalitäten in einem Wissensgebiet [2]. Das Konzept der Ontologie entstammt der Philosophie und wurde in die Informatik übernommen.

"Since ancient times, ontology, the analysis and categorisation of what exists, has been fundamental to philosophical enquiry. But, until recently, ontology has been seen as an abstract, purely theoretical discipline, far removed from the practical applications of science. However, with the increasing use of sophisticated computerised information systems, solving problems of an ontological nature is now key to the effective use of technologies supporting a wide range of human activities.[...]"[3]

Sie können allgemeines Wissen über die Struktur von Informationen definieren und spezielles Fachwissen modellieren. Dazu wird das Wissen als Netzwerk aus hierarchischen Klassen mit Beziehungen untereinander organisiert. Ontologien werden benutzt, um Wissen im Computer abzubilden und dienen als gemeinsame Wissensbasis zwischen Experten und zur Weitergabe von Wissensbeständen.

Solange alle Parameter bekannt sind und keine Unsicherheiten bestehen lassen sich logische Schlüsse aus dem repräsentierten Wissen ziehen.

Bestandteile von Ontologien sind:

- Klassen (oder Begriffe) dienen der Beschreibung gemeinsamer Eigenschaften verschiedener Objekte. Siehe Abbildung 2 "Gebäude" und "Tür".
- Instanzen von Klassen sind real existierende Objekte. So ist "Karlsruhe" eine Instanz von "Stadt".
- Relationen bilden die Beziehungen zwischen bestehenden Instanzen ab. Sie werden auch als Eigenschaften oder Attribute bezeichnet und als Pfeil dargestellt. Das KIT "liegt in" Karlsruhe ist eine Relation. Einstellige Relationen werden Attribute genannt ("Höhe" in Abbildung 2).
- Vererbungen: Relationen von Klassen können vererbt werden und ermöglichen spezialisierte Klassen. In Abbildung 2 erben "Haus" und "Garage" von "Gebäude".
- Axiome sind Aussagen, die innerhalb der Ontologie immer wahr sind. Sie dienen der Definition von Wissen welches sich nicht aus den Klassen und Relationen ableiten lässt.

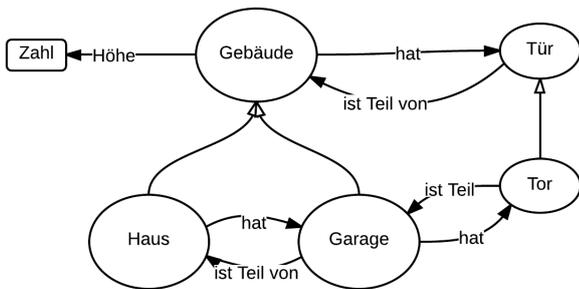


Abbildung 2. Eine einfache Ontologie über den Zusammenhang zwischen Haus und Garage

B. Bayessche Netze

Mit Bayesschen Netzen lassen sich bedingte Wahrscheinlichkeiten effektiv modellieren, darstellen und auswerten. Es besteht aus Knoten die jeweils eine Zufallsvariable darstellen, und Kanten die die zugehörige bedingte Abhängigkeit zwischen den Variablen angeben. Der Knoten, von dem eine Kante ausgeht nennt sich Elternknoten. In Abbildung 3 ist "Regen" ein Elternknoten von "Sprinkler". "Sprinkler" besitzt somit eine bedingte Wahrscheinlichkeitsverteilung. Diese bedingte Wahrscheinlichkeitsverteilung muss für jede eingehende Kante eines Knotens gegeben sein ("Sprinkler", "Gras nass?"). Hat ein Knoten keine eingehenden Kanten ("Regen"), so handelt es sich bei seiner Wahrscheinlichkeitsverteilung um eine unbedingte Verteilung. In Abbildung 3 sind die Zufallsvariablen "Regen (R), Sprinkler (S) und Gras nass? (G)" als Ovale gezeichnet. Die Tabellen geben die Wahrscheinlichkeitsverteilungen der zugehörigen Zufallsvariablen an und die Pfeile symbolisieren die Abhängigkeiten. So hängt die Zufallsvariable "Sprinkler" von "Regen" ab.

Die gemeinsame Verteilung von X_1, \dots, X_n berechnet sich durch:

$$P(X_1, \dots, X_n) = \prod_{i=1}^n (P(X_i | \text{Eltern}(X_i)))$$

Angewendet auf das Bayessche Netz in Abbildung 3:

$$P(G, S, R) = \underbrace{P(G|S, R)}_{i=1} * \underbrace{P(S|R)}_{i=2} * \underbrace{P(R)}_{i=3}$$

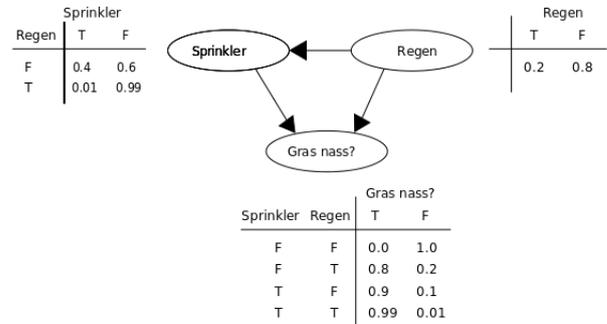


Abbildung 3. Ein einfaches Bayessches Netz (nach: <http://en.wikipedia.org/wiki/File:SimpleBayesNet.svg>)

Die unbedingte Wahrscheinlichkeit für "Der Sprinkler ist an" = $S = T$ berechnet sich:

$$P(S = T) = P(S|R) * P(R) + P(S|\bar{R}) * P(\bar{R}) = 0,01 * 0,2 + 0,4 * 0,8 = 0,32$$

D.h. der Sprinkler ist mit einer Wahrscheinlichkeit von etwa 32% an. Die Berechnung mit welcher Wahrscheinlichkeit $P(S = T)$ also das Gras nass ist, ist etwas länger:

$$\begin{aligned}
 P(G = T) &= \sum_{R,S \in \{T,F\}} P(G = T, R, S) \\
 &= \sum_{R,S \in \{T,F\}} P(G = T, R, S) P(S|R) P(R) \\
 &= (0,99 * 0,01 * 0,2) + (0,8 * 0,99 * 0,2) \\
 &\quad + (0,9 * 0,4 * 0,8) + (0 * 0,6 * 0,8) \\
 &= 0,448
 \end{aligned}$$

Im Gegensatz zu Ontologien beruhen Bayessche Netze aber auf einer festen Konfiguration der Zufallsvariablen und deren Beziehungen zueinander [4]. Sie kennen keine Klassen aus denen sich Objekte, in verschiedenen Konfigurationen, ableiten lassen. Sollte nun ein weiterer Sprinkler hinzukommen, müsste ein neues Bayessches Netz definiert werden. Daher können sie nicht dazu benutzt werden prinzipielle Zusammenhänge ähnlicher Objekte abzubilden und diese auf verschiedene Situationen anzuwenden [5].

III. VORSTELLUNG VERSCHIEDENER ANSÄTZE ZUR AUSNUTZUNG VON MODELLWISSEN BEIM PROBABILISTISCHEM SCHLUSSFOLGERN

A. Objektorientiertes Weltmodell nach [6]

Bei der Objektorientierten Weltmodellierung wird das Modellwissen ähnlich wie mit Ontologien modelliert. So besitzen

Klassen Attribute mit einer bedingten Wahrscheinlichkeit und sind über Relationen vernetzt. Diese Klassen sollten durch Vererbungen immer weiter verfeinert werden können, denn sie dienen später im Weltmodell als Vorlage für konkrete Objekte. Diese Objekte dienen der hierarchischen Repräsentation des Wissens und werden je nach Wissensgrad (Sensorwerte) und Anforderung verfeinert. Sie bilden eine Abstraktionspyramide, wie sie in Abbildung 4 zu sehen ist.

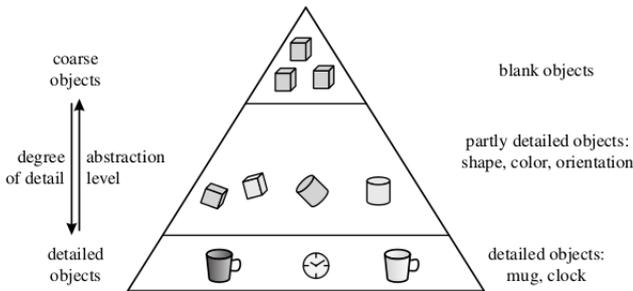


Abbildung 4. Abstraktion von Objekten in einer Szene (aus: [6])

Eine Instanz im Weltmodell in Abbildung 5 kann z.B. von einem unbeweglichen Objekt zu einer Ampel und von einem beweglichen zu einem Menschen verfeinert werden. Die Unbeweglichkeit der Ampel ist für die Situationserkennung eigentlich unwichtig, vereinfacht aber die Erkennung, da genauere oder weitere Sensorwerte über den Abstraktionsgrad direkt ins Weltmodell eingepflegt werden können.

Der Abstraktionsgrad hängt von den bekannten Attributen und nicht von deren Unsicherheitsfaktor ab. Das Weltmodell wird jedoch nur soweit wie nötig verfeinert. Nicht relevante Informationen werden ausgeblendet. Soll ein Weg berechnet werden so reicht die Information über die Größe und Position der Objekte. Der Typ ist irrelevant. Soll jedoch erkannt werden, ob an einer konkreten roten Ampel gehalten werden muss, muss eine genauere Sicht ausgewertet werden. Es ist nicht nur wichtig zu erkennen, dass die Ampel rot ist, sondern noch der Zusammenhang zur konkreten Fahrsituation geschlussfolgert werden.

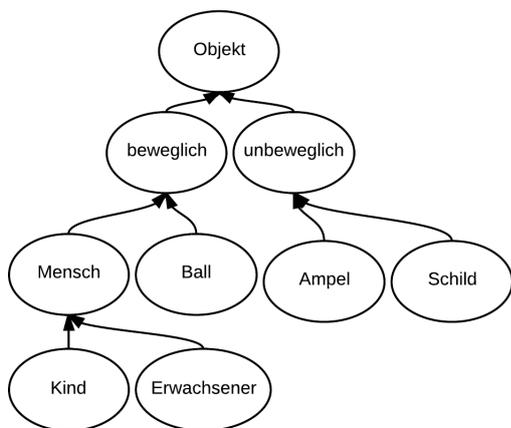


Abbildung 5. Beispiel der Verfeinerung im Weltmodell (nach: [6])

keitsverteilung, sowie das Auflösen von Konflikten findet in einer konkreten Instanz eines Objektes statt. Jede Instanz besitzt darüber hinaus einen Unsicherheitsfaktor, der angibt wie wahrscheinlich sie existiert und mit welcher Wahrscheinlichkeit ihre Zuordnung zu einer Klasse richtig ist. Dieser Mechanismus beruht größtenteils auf dem vorhandenen Modellwissen und bedingten Wahrscheinlichkeiten. Mit Zeitstempeln lassen sich die Instanzen im Weltmodell als Szenen interpretieren um komplexere Zusammenhänge wie Bewegungen darstellen und erkennen zu können. Eine Konsistenzprüfung sollte aus Performance Gründen zyklisch stattfinden. Sie stellt sicher, dass das Modell konsistent und valide ist. Außerdem sollten nicht relevante Informationen aus dem Modell entfernt werden. Um eine Szene aufzubauen und möglicherweise widersprüchliche Sensorwerte auszugleichen besitzt jedes Instanz zusätzlich eine Ausbreitungsunsicherheit. Sie drückt aus wie wahrscheinlich es ist, dass die jeweilige Instanz verschwindet und beeinflusst das Vertrauen in die Existenz der Instanz über die Zeit. So ist es wenig wahrscheinlich, dass die Straße plötzlich verschwindet - sie besitzt eine große Lebensdauer. Die Farbe einer Ampel kann sich jedoch sehr schnell ändern und besitzt daher eine kurze Lebensdauer. Wird die Existenz einer Instanz über längere Zeit von keinem Sensor bestätigt, so wird sie aus der Szene entfernt. Das Weltmodell wird aus vordefiniertem Wissen und Sensorinformationen gebildet, wobei vordefiniertes Wissen, z.B. das Vorhandensein von Fußgängern, solange es noch nicht durch die passenden Sensorinformationen bestätigt wurde, einen hohen Unsicherheitsfaktor besitzt. Erkennt nun ein Sensor ein Objekt und es kann noch keiner Instanz im Weltmodell zugeordnet werde, so wird es als neue Instanz mit höchstmöglichem Abstraktionsgrad hinzugefügt. Über eine Relation kann diese neue Instanz auch ein Attribut einer schon im Weltmodell bestehenden sein kann. Der große Unterschied zu Multi Entity Bayesian Networks ist, dass das Weltmodell kontinuierlich angepasst und nicht jedes mal neu kompiliert wird.

B. Multi Entity Bayesian Networks

Multi Entity Bayesian Networks (MEBN) sind die Erweiterung Bayesscher Netze um eine Art Klassenlogik. Nach [4] und [8] lassen sich mit MEBN wie mit Ontologien prädikatenlogische Aussagen formulieren. Das Hintergrundwissen über konkrete Attribute und Beziehungen wird in MEBN-Fragmenten (MFrag) gespeichert. Diese wiederum sind zu MEBN-Theorien (MTheories) gruppiert. Sie sind also ähnlich zu den Klassen in Ontologien. Ein MFrag kann als Gerüst zum Bilden beliebig vieler situationsspezifischer Bayesscher Netze gesehen werden. Es wird instanziiert, indem seine Argumente an Werte gebunden werden. Ein MFrag enthält Zufallsvariablen und einen Graphen, der die Abhängigkeiten zwischen ihnen definiert. Es gibt drei Arten solcher Zufallsvariablen:

- Ein Kontextknoten definiert Randbedingungen, die eingehalten werden müssen.
- Ein Eingabeknoten kann Zufallsvariablen außerhalb seines MFrag beeinflussen oder von ihnen beeinflusst werden, seine Wahrscheinlichkeitsverteilung wird aber im MFrag definiert.

Konsistenzprüfung, die Auswertung der Wahrscheinlich-

- Die Wahrscheinlichkeitsverteilung innerer Knoten (resident nodes) ist nur von deren Elternknoten im MFrage abhängig.

Eine Menge von MFrage repräsentiert die mehrdimensionale Verteilung der Instanzen ihrer Zufallsvariablen. Eine MTheorie ist eine Menge von MFrage, die zusammen Konsistenzbedingungen einhalten und sicherstellen, dass eine eindeutige mehrdimensionale Verteilung gegeben ist. Laut [8] sind MEBN eine gute Möglichkeit gleichartige Strukturen in Bayesschen Netzen abzubilden. Auf den aus ihnen generierten Instanzen lassen probabilistische Schlussfolgerungen über konkrete Situationen ziehen und auch die MTheorien mit Hilfe gegebener Situationen verbessern.

C. Probabilistic-OWL (PR-OWL)

PR-OWL ist eine Erweiterung der Web Ontology Language (OWL), um probabilistische Ontologien erstellen zu können [9]. Mit Hilfe von MEBN eignet sich PR-OWL zur konsistenten Abbildung und Auswertung von unsicherem Wissen, welches auch ausgetauscht und weiterverwendet werden kann [8]. Mit PR-OWL lassen sich Probabilistische Ontologien bilden und mit verschiedenen Programmen auswerten. Eine Probabilistische Ontologie besteht aus mindestens einer MTheorie, was in PR-OWL die Objekteigenschaft "hat MTheorie" innerhalb eine Menge von MFrage ist. Ein individueller MFrage besteht wiederum aus Knoten, die dann eine Zufallsvariable darstellen. Jede Zufallsvariable besitzt mögliche Konfigurationen der Klasse "Entity", die durch die Eigenschaft "hasPossibleValues" verknüpft sind. Die Wahrscheinlichkeitsverteilung wird durch die Klasse "ProbabilityDistribution" repräsentiert und an die einzelnen Knoten gebunden. Dies ist in Abbildung 7 vereinfacht dargestellt. Eine Probabilistische Ontologie lässt sich mit einem OWL kompatiblen Editor wie Protégé erstellen. D diese nicht darauf ausgelegt sind MFrage zu definieren sind sie nicht gut dazu geeignet. Einfacher geht es mit UnBBayes, welches auch ein Tool zur Auswertung von MEBN integriert. Zum Auswerten muss eine so erstellte Probabilistische Ontologie mit Werten befüllt werden. Dann kann ein der konkreten Situation entsprechendes Bayessches Netz daraus kompiliert werden.

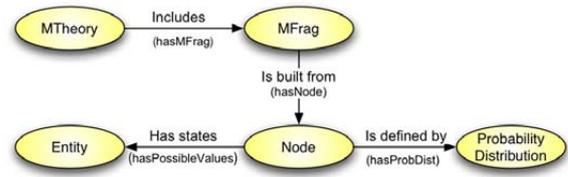


Abbildung 7. Vereinfachte Übersicht über PR-OWL (aus: [8])

D. Object-Oriented Probabilistic Relational Language

Die Komplexität von PR-OWL macht die Generierung des MEBN aufwändig. Für viele Anwendungszwecke reicht ein einfacheres Relationales Modell, aus dem dann Bayessche Netze generiert werden. Dies bewerten [5] und [10] auf welche der folgende Abschnitt basiert. Die Informationen zur Bildung des Modellwissens können zum größten Teil aus einer relationalen Datenbank gebildet werden. Da die verwendete relationale Datenbank keine Inverse¹ zulässt, besitzt das Hintergrundwissen keine Inverse und es können nur abzählbar viele mögliche Attributwerte existieren, die im vorab definiert werden müssen. In Abbildung 8 steht eine konkrete Instanz von "Haus" in Relation zu möglicherweise mehreren Instanzen von "Fenster" und "Tür". Ein konkretes Haus kann also mehrere Türen und Fenster besitzen. Nach [7] definiert die Object-Oriented Probabilistic Relational Language (OPRL) ein Entity-Relationship-Modell (ER-Modell) aus zwei Komponenten welches als Gerüst zum Bilden konkreter Instanzen dient. Die relationale Komponente definiert die Abhängigkeiten zwischen Klassen in einer Domäne, während der probabilistische Teil die bedingten Wahrscheinlichkeiten angibt. Die Klasseninformationen können automatisiert aus einer Relationalen Datenbank ausgelesen werden, während der probabilistische Teil festgelegt werden muss. Dieser definiert die bedingten Wahrscheinlichkeiten einzelner Attribute.

Das Weltmodell wird gebildet, indem Sensorinformationen als Instanzen in das Weltmodell eingefügt werden. Diese Instanzen werden mit den als Relationen hinterlegten Vorlagen

¹Eine Inverse ist eine beidseitige Relation. Um Inverse auszuschließen, muss die Relation gerichtet sein.

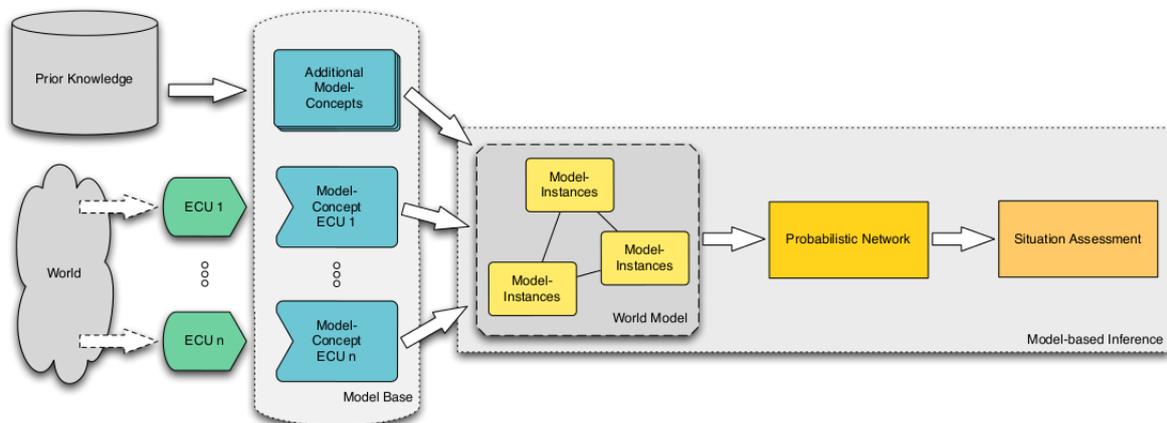


Abbildung 6. Informationsfluss beim Schlussfolgern mit der Object-Oriented Probabilistic Relational Language (aus: [7])

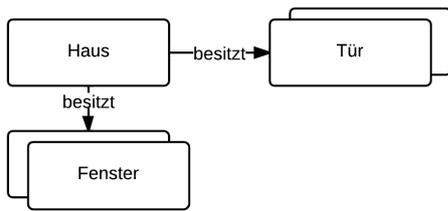


Abbildung 8. Gerüst einer relationalen Datenbank

abgeglichen, die definierten Abhängigkeiten eingepflegt und laufend aktualisiert. Wie aus Abbildung 6 ersichtlich, liegt ein entscheidender Unterschied zum Objektorientierten Weltmodell darin, dass die Generierung eines Bayesschen Netzes zur Auswertung - es kann auch ein Bayessches Netz pro Objekt gebildet werden - von der Bildung des Weltmodells entkoppelt ist und die beschreibende Sprache für das Modellwissen weniger mächtig ist.

IV. ZUSAMMENFASSUNG UND AUSBLICK

Das Objektorientierte Weltmodell nutzt das Hintergrundwissen um aus Sensorinformationen Instanzen von Klassen zu bilden, Zusammenhänge zwischen ihnen darzustellen und kontinuierlich ins Weltmodell einzupflegen. Dieses lässt sich dann zum probabilistische Schlussfolgern nutzen.

Multi Entity Bayesian Networks erweitern Bayessche Netze um eine Art Klassenlogik und ermöglichen so die Abbildung allgemeiner Zusammenhänge zwischen ähnlichen Zufallsvariablen. Mit diesem Wissen lassen sich konkrete Instanzen zum probabilistischen Schlussfolgern generieren.

Probabilistic-OWL baut auf MEBN auf und erweitert OWL um unsicheres Wissen. Mit OWL und MEBN lässt sich Hintergrundwissen zum probabilistischen Schlussfolgern nutzen.

Die Object-Oriented Probabilistic Relational Language beschränkt sich bei der Modellierung des Hintergrundwissens auf relationale Schemata. Das Hintergrundwissen wird im Weltmodell genutzt um die Attribute und Relationen konkreter Instanzen mit Unsicherheiten zu versehen.

Zur weiteren Forschung bietet sich der genauere Vergleich dieser Ansätze im Bezug zur Anwendbarkeit auf konkrete Fragestellungen und der Geschwindigkeit der Auswertung an.

LITERATUR

- [1] A. Kuwertz and G. Schneider, "Ontology-Based Meta Model in Object-Oriented World Modeling for Interoperable Information Access (to appear)," in *Proceedings of the Eighth International Conference on Systems (ICONS)*, Jan. 2013.
- [2] N. F. Noy and D. L. McGuinness, "Ontology development 101: A guide to creating your first ontology." CTAN: http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html.
- [3] F. C. Bennett B., "Formal Ontology in Information Systems," in *Proceedings of the Fourth International Conference (FOIS 2006)*, 2006.
- [4] Kathryn B. Laskey, "MEBN: A language for First-Order Bayesian Knowledge Bases. Artificial Intelligence," 2007.
- [5] Lise Getoor, Nir Friedman, Daphne Koller, Avi Pfeffer and Ben Taskar, *Introduction to Statistical Relational Learning*, ch. 5 Probabilistic Relational Models. MIT Press, 2007.
- [6] Gheta, I.; Heizmann, M.; Beyerer, J., "Object oriented environment model for autonomous systems," in *SWIFT 2008, Second Skövde Workshop on Information Fusion Topics. Proceedings*, 2008.
- [7] Thomas Schamm, J. Marius Zöllner, "A Model-based Approach to Probabilistic Situation Assessment for Driver Assistance Systems," 2011.
- [8] Rommel Carvalho, Kathryn Laskey, Paulo Costa, Marcelo Ladeira, Laécio Santos and Shou Matsumoto, *UnBBayes: Modeling Uncertainty for Plausible Reasoning in the Semantic Web*, ch. 1, pp. 1 – 28. In-Tech, 2010.
- [9] K. J. L. Paulo Cesar G. da Costa, Kathryn B. Laskey, "PR-OWL: A Bayesian Ontology Language for the Semantic Web," in *Uncertainty Reasoning for the Semantic Web I Lecture Notes in Computer Science Volume 5327*.
- [10] Catherine Howard and Markus Stumptner, "Applying OPRMs to Recursive Probability Models," *DOI:European Conference on Artificial Intelligence ECAI*, p. 2, 2006.