Diplomarbeit

Modellierung eines intensivmedizinischen Behandlungsprotokolls zur Validierung anhand realer Patientendaten

Martin Scholz

Diplomarbeit
am Fachbereich Informatik
der Universität Dortmund

Dortmund, 18. Juli 2001

Betreuer:

Prof. Dr. Katharina Morik
Dipl.-Inform. Oliver Ritthoff

Danksagung

An dieser Stelle möchte mich ganz herzlich bei meinen beiden Betreuern Katharina Morik und Oliver Ritthoff für deren Engagement, hilfreiche Anregungen und konstruktive Kritik bedanken.
Weiterhin bedanken möchte ich mich bei Herrn Imhoff, für die Bereitstellung der Datensätze und besonders für die unkomplizierte fachliche Unterstützung bei allen Fragen rund um die Intensivmedizin.
Den Mitarbeitern, Diplomanden und studentischen Hilfskräften des Lehrstuhls Informatik VIII sei für die Hilfe bei zahlreichen technischen Problemen und das motivierende Arbeitsumfeld gedankt.
Abschließend möchte ich mich bei meiner Freundin Sandra für das Verständnis der letzten Wochen bedanken, in denen die Vorfreude auf unseren Nachwuchs manchmal recht kurz gekommen ist.

Inhaltsverzeichnis

Abbildungsverzeichnis

Tabellenverzeichnis

Einleitung

Die Betreuung von Patienten auf modernen Intensivstationen ist eine verantwortungsvolle und schwierige Aufgabe. Ärzte verlassen sich bei der Wahl und Durchführung von Interventionen auf Fachwissen und Berufserfahrung. Eine zunehmend genaue und umfangreiche Erfassung von Patientenvariablen erschwert nunmehr die Identifikation entscheidungsrelevanter Informationen. Im klinischen Alltag treten unweigerlich vermeidbare Fehler auf. Für diese Arbeit sind informationsbedingte Kunstfehler von besonderem Interesse. Entscheidungshilfe-Systeme können helfen, diese Art von Fehlern zu vermeiden. Dazu sollten sie in der Lage sein, die relevanten Aspekte erhobener Patientenvariablen zu erkennen und automatisch auszuwerten. Sie gehen über die weit verbreiteten Alarm-Systeme, die kritische Patientenzustände signalisieren, hinaus, da sie in der Lage sind, medizinische Interventionen vorzuschlagen. Neben einer aufwendigen sukzessiven Verfeinerung von Behandlungsprotokollen, gestützt auf fortlaufende klinische Validierungen, ist eine Entwicklung und Validierung solcher Protokolle am Rechner denkbar. Ziel dieser Diplomarbeit ist, die Durchführbarkeit dieses Ansatzes zu testen. Protokollierte klinische Verlaufsdaten sollen dabei helfen, unangemessene Protokollteile zu identifizieren.
Für diese Arbeit steht ein intensivmedizinisches Behandlungsprotokoll zur Verfügung, das exemplarisch validiert werden soll. Gegenstand des Protokolls sind Teilbereiche der Intensivmedizin, besonders das hämodynamische System und der Flüssigkeitshaushalt. Da das Protokoll in einer Papierversion als Flussdiagramm vorliegt, besteht ein Großteil dieser Arbeit in einer geeigneten Operationalisierung. Um die Diskutierbarkeit des umgesetzten Modells zu gewährleisten und Modell-Revisionen zu erleichtern, wird eine deklarative Umsetzung in einer logischen Repräsentationssprache vorgenommen.
Schließlich werden einige mit der Wissensbasis durchgeführte Experimente beschrieben. Zur Validierung stehen protokollierte Behandlungsdaten zur Verfügung, aus denen neben den erhobenen Patientenvariablen auch die jeweils durchgeführten Interventionen ersichtlich sind. Nach einer geeigneten Ergänzung der Wissensbasis wird diese anhand der Daten validiert.

In Kapitel 1 wird zunächst der Hintergrund der Diplomarbeit erläutert. Nachdem Art und Umfang vermeidbarer Fehler in der Medizin beschrieben wurden, wird auf die Bedeutung standardisierter Behandlungsprotokolle und Entwicklungsalternativen zu deren Aufbau eingegangen. Weiterhin werden benötigte medizinische Grundbegriffe und schließlich einige verwandte Arbeiten vorgestellt.
In Kapitel 2 werden die Ziele dieser Arbeit beschrieben. Zuvor eher vom medizinischen Standpunkt aus betrachtete Problemstellungen werden nun aus der Sicht der Informatik konkretisiert. Dabei werden die verwendeten Methoden motiviert.
Einen Schwerpunkt stellt die in Kapitel 3 beschriebene Übersetzung des Flussdiagramms in ein wissensbasiertes System dar. Nach einer einführenden Beschreibung des umzusetzenden Protokolls wird zunächst allgemein auf wissensbasierte Systeme eingegangen. Schließlich wird die Wahl einer konkreten Architektur für die Repräsentation des Wissens

hergeleitet, anhand der das umgesetzte Modell beschrieben wird.
Die Validierung des Modells ist Gegenstand des Kapitels 4. Ausgehend von einer Beschreibung der verwendeten Datensätze wird der Ausbau des Modells zu einem Kritik-System beschrieben, das die Therapievorschläge des Protokolls den Interventionen aus den Verlaufsdaten gegenüberstellt. Abschließend wird ein weiteres Experiment vorgestellt, mit dessen Hilfe die Stetigkeit gemachter Protokollvorschläge im zeitlichen Verlauf untersucht wird.
Kapitel 5 fasst die Ergebnisse der Arbeit zusammen und gibt einige Anregungen für weiterführende Arbeiten.

1. Hintergrund der Diplomarbeit

1.1. Fehler in der Medizin

Laut eines Berichts des „Institute of Medicine“ aus den USA [KCD00] befinden sich, selbst nach vorsichtigen Schätzungen, vermeidbare medizinische Fehler unter den acht häufigsten Todesursachen überhaupt, noch vor Verkehrsunfällen, Brustkrebs und AIDS. Grundlage dieser Schätzungen sind zwei große amerikanische Studien, bei der die Quote nachgewiesener medizinischer Fehler pro Krankenhausaufenthalt zwischen 2.9 und 3.7 Prozent lagen, von denen wiederum zwischen 6.6 und 13.6 Prozent zum Tode führten.

Eine sehr ernstzunehmende Fehlerquelle ist die Verabreichung von Medikamenten. So sterben jährlich in den USA etwa 7000 Menschen an den Folgen einer fehlerhaften Medikation, während zum Vergleich „nur“ 6000 Menschen durch einen Unfall am Arbeitsplatz versterben. Medikationsfehler treten dabei in besonderem Maße in Krankenhäusern auf. Zu beachten ist, dass nicht alle Fehler vermeidbar sind.

Typische vermeidbare Fehler sind die Verwechslung von Medikamenten am Krankenbett, die Verschreibung einer ungeeigneten Dosis, die Verabreichung einer anderen, als der verordneten Dosis, die Verordnung eines für den beabsichtigten Zweck ungeeigneten Medikaments, oder auch eine medikamentöse Therapie zur Verfolgung eines unangemessenen Zieles. Auch wenn unter mehreren möglichen Medikamenten ausgerechnet eines verordnet wird, auf das der Patient bekanntermaßen allergisch reagiert, etwa aus Nachlässigkeit des behandelnden Arztes, oder weil die Information nicht am Krankenbett verfügbar ist, so ist dies ein vermeidbarer Fehler. Ein Beispiel für einen unvermeidbaren Fehler wäre die Verordnung eines Medikaments, dessen Unverträglichkeit bislang nicht festgestellt wurde.

Die Gefahr ärztlicher Fehler ist besonders im Umfeld intensivmedizinischer Betreuung und nach Operationen gegeben. Dies lässt sich einer im Bericht des Institute of Medicine aufgeführten Studie entnehmen, die die Behandlung 1047 entsprechender Patienten in einem großen amerikanischen Lehrkrankenhaus untersucht. Dabei erfuhren 480 der betrachteten Patienten eine vermeidbar fehlerhafte Behandlung. Bei immerhin 185 dieser Patienten führte diese unangemessene Entscheidung zu ernsthaften gesundheitlichen Schäden, oder sogar zum Tode. Der Anteil fehlerhaft behandelter Patienten liegt damit bei 45.8 Prozent, der Anteil der Patienten, für die dies ernsthafte Konsequenzen hatte, bei 17.7 Prozent. Die Wahrscheinlichkeit, einer fehlerhafte Entscheidung zum Opfer zu fallen, stieg mit der Aufenthaltsdauer, etwa um 6 Prozent pro Tag.

Neben den tragischen menschlichen Einzelschicksalen, entstehen bei einer so hohen Zahl folgenschwerer medizinischer Fehler bereits ernstzunehmende volkswirtschaftliche Schäden. So werden die rein finanziellen Folgeschäden durch vermeidbare medizinische Fehler in den USA jährlich auf einen Betrag zwischen 17 und 29 Milliarden US-Dollar geschätzt. Anders interpretiert stehen Bestrebungen zur Erhöhung der Patientensicherheit keine finanziellen Hürden im Wege, sondern diese führen, volkswirtschaftlich gesehen, eher noch zur Abschöpfung von Einsparungspotentialen.

Expertenmeinungen zufolge ist das Problem medizinischer Fehler kein rein amerikanisches Problem. In Europa ist die Situation sogar eher noch ernster. Als eine der Hauptursachen für Fehler ist Informationsüberflutung im klinischen Umfeld zu nennen [Imh00].

Will man dem Problem wirkungsvoll begegnen, so ist es notwendig, die fehleranfälligen Teilaspekte in der Patientenversorgung zu finden und zu beseitigen. Ein einfaches Beispiel ist, darauf hinzuarbeiten, dass häufig verwechselte Medikamente besser (z.B. optisch, andere Namensgebung) unterscheidbar werden. Die Identifikation und Sanktionierung von Schuldigen in jedem Einzelfall vermag hingegen nur einen geringen Beitrag zur Qualitätssicherung des Gesamtsystems zu leisten. Dies ist in gewisser Weise vergleichbar mit dem Straßenverkehr, wo aufgrund von Unfallstatistiken bestimmte Streckenabschnitte mit Geschwindigkeitsbegrenzungen, Überholverboten und Warnhinweisen versehen werden. Hier wird versucht gezielt den Einflüssen, die in der Vergangenheit häufig zu Unfällen geführt haben, entgegen zu wirken. Geeignete Statistiken sind hier ein zur Steigerung der Sicherheit wertvoller Anhaltspunkt.

Das Institute of Medicine plädiert in diesem Sinne in seinem Bericht [KCD00] für einen Rahmen, in dem Fehlern besser nachgegangen werden kann, einerseits durch die Pflicht, gewisse Vorfälle zu melden, andererseits durch die Ermutigung zur freiwilligen und unsanktionierten (z.B. anonymen) Veröffentlichung von Fehlern. Bei der systematischen Fehlersuche und nachfolgenden Analyse kann moderne informationsverarbeitende Technologie hilfreich sein. Elektronische Patientenakten, also die elektronische Dokumentation erfasster Verlaufsdaten, medizinischer Interventionen etc., ist dabei ein wichtiger Schritt, um der technischen Unterstützung den Weg zu bereiten. Erklärtes Ziel ist eine spürbare Verbesserung der Patientensicherheit. Die Erstellung von Richtlinien, die helfen, bekannt gewordene Sicherheitsrisiken auszumerzen, ist ein wichtiges Mittel hierfür. Durch den Einsatz moderner informationsverarbeitender Technologien sind jedoch noch weitaus tiefgreifendere Qualitätssicherungsmaßnahmen möglich.

Eine der bereits in dem Bericht gemachten Empfehlungen ist die Verwendung von Entscheidungshilfe-Systemen, beispielsweise einer pharmazeutischen Wissensbasis, die Medikamentenverordnungen gegenprüft, unter anderem um typische Fehler bei der Medikation zu vermeiden. Mit einem solchen System können Medikationen, die zu allergischen Reaktionen, problematischen Wechselwirkungen mit anderen Medikamenten oder gar zu Überdosierungen führen, frühzeitig erkannt werden. Ein wissensbasiertes System, das neben der Unterstützung bei der Arzneimittelverordnung diese auch automatisch dokumentiert, und darüber hinaus eine Erkennung von Elektrolytentgleisungen, Nierenfunktionsverschlechterung und Medikamentenüberdosierungen leistet, wird in [Wai98] vorgestellt. Das System ist in das Gießener Krankenhausinformationssystem integriert und meldet seit 1996 routinemäßig Monitoring-Warnhinweise direkt auf die Station.

Darüber hinaus bietet die Informatik aber noch weitere Möglichkeiten zur Qualitätssicherung in der Medizin. Diese Diplomarbeit beschäftigt sich mit einem rechnerbasierten, intensivmedizinischen Behandlungsprotokoll, als einen Ansatz, Fehlern entgegenzuwirken. Die Entscheidungsfindung verantwortlicher Mediziner wird durch solche Unterstützungssysteme vereinfacht. Weiterhin stehen getroffene Entscheidungen automatisch einer zweiten „Meinung“, dem Vorschlag des Systems gegenüber. Im nächsten Abschnitt wird zunächst etwas allgemeiner auf den Einsatz von Behandlungsprotokollen in der Medizin eingegangen.

1.2. Behandlungsprotokolle in der Medizin

Ein wirkungsvolles Mittel, um die Sicherheit von Patienten zu steigern und die Behandlungsqualität zu erhöhen, ist die Erstellung und Verwendung geeigneter rechnerbasierter Unterstützungssysteme. Mit Hilfe solcher Systeme kann die Zahl menschlicher Fehler (siehe Abschnitt 1.1) reduziert werden, beispielsweise indem die jeweils aktuell verordnete Medikation auf typische Fehler hin geprüft wird, oder indem Patientenvariablen automatisch überwacht werden, um auf übersehene kritische Zustände aufmerksam zu machen. In diesem Abschnitt wird etwas ausführlicher auf regelbasierte Behandlungsprotokolle für die Medizin eingegangen. Diese helfen nicht nur dabei, menschliche Fehler zu bemerken, sondern sie können darüber hinaus Interventionen vorschlagen.
In [Mor98] stellt A. Morris ein Beispiel für ein derartiges Behandlungsprotokoll vor. Es handelt sich dabei um ein Beatmungsprotokoll für das Krankheitsbild des akuten Lungenversagens. Die jeweils beste Behandlung bei diesem Krankheitsbild zu finden, ist eine äußerst anspruchsvolle intensivmedizinische Aufgabe. Die Entwicklung des Protokolls am LDS Hospital, Salt Lake City, Utah, USA, nahm 25 Personen-Jahre in Anspruch. Mittlerweile ist das Protokoll als marktreifes Produkt erhältlich. Einige weitere Systeme sind in [Mik94] aufgezählt.
Morris bezeichnet in [Mor98] eine über Protokolle standardisierte Behandlung von Patienten als einen Paradigmenwechsel in der Medizin. Die bisherige Auffassung ist die, dass jeder Arzt aufgrund der Ausbildung und persönlichen Erfahrungen für seine Patienten in jeder Situation die beste Therapie zu finden vermag. Richtet sich die Behandlung jedoch nach standardisierten Behandlungsprotokollen, so wird die Therapie jeweils aus den Anweisungen des Protokolls abgeleitet. Das Protokoll soll dabei derart beschaffen sein, dass es nachgewiesenermaßen in allen Fällen eine wirkungsvolle, möglichst sogar die beste Therapie vorschlägt. Für den Nachweis der Wirksamkeit einer Therapie, oder der Überlegenheit einer Therapie gegenüber einer anderen, ist die Validierung in Form von klinischen Studien notwendig.
Als Vorteil solcher Behandlungsprotokolle nennt Morris besonders die Vorzüge einer Standardisierung, sowohl der Behandlung, als auch der verwendeten medizinischen Begriffe.
Er zeigt an Beispielen, dass medizinische Fachbegriffe oft nicht einheitlich verwendet werden. Dies kann zu Problemen führen, etwa wenn verschiedene Intensivmediziner die 236 Variablenkategorien, die bei der Behandlung von akutem Lungenversagen einfließen, aufeinander abbilden müssen. Soll versucht werden, die Vorteile einer Therapie gegenüber einer anderen durch klinische Studien aufzuzeigen, so ist es von größter Bedeutung, dass die Begriffe und Anweisungen präzise definiert, und die Methoden der Studie ebenso präzise reproduzierbar sind, da andernfalls keine wissenschaftlich verwertbaren Aussagen möglich sind. Die Explizierung der Begriffe und Interventionen ermöglicht also oft überhaupt erst die Vergleichbarkeit der Behandlung über regionale Grenzen hinweg. Hierfür schlägt Morris scharfe Grenzwerte im Falle numerischer Daten vor, und plädiert für eine logische Repräsentation von Begriffen und Regeln.
Zur Rolle der Validierung der Wirksamkeit von Therapien merkt Morris an, dass es für viele medizinische Interventionen keine Daten gibt, die zuverlässig belegen, dass sie mehr nutzen als schaden. Ein Problem bei der Validierung ist die Tatsache, dass der Nutzen einer Therapie oft verhältnismäßig gering ist, und dass daher systematische Fehler bei der Erhebung und unkontrollierte Faktoren, die den Patientenzustand beeinflussen, statistisch nur schwer von der eigentlichen Zielgröße, dem Einfluss der Therapie separierbar sind.

[Ohm96] gibt eine Übersicht über typische Störgrößen bei der Validierung medizinischer wissensbasierter Systeme, und verschiedenen Studientypen, mit ihren diesbezüglichen Vor- und Nachteilen. Zu den Störgrößen, die Ergebnisse einer klinischen Studie fälschlicherweise verbessern können, zählt beispielsweise Übereifer, wenn Systeme von ihren Entwicklern, oder zumindest am Ort der Entwicklung validiert werden. Ein anderes Problem stellt der sogenannte „Hawthorne-Effekt“ dar. Dieser bezeichnet eine verbesserte Leistung des beteiligten medizinischen Personals aufgrund der Beobachtung im Rahmen von Studien. Bei Studien, die auf Vergleichen mit historischen Kontrollgruppen basieren, sind die Einflüsse sekularer Trends zu beachten, das heißt es darf nicht außer Acht gelassen werden, dass sich die Medizin generell im Laufe der Zeit verbessert hat, beispielsweise durch den Einsatz zuverlässigerer Technologien. Insbesondere die Eingrenzung dieser und weiterer Störgrößen machen Studien aufwendig und kostspielig. Grundsätzlich gilt aber, dass mit zunehmender Stichprobengröße auch die statistische Separation zunehmend besser gelingt, den Ergebnissen also größere Signifikanz verliehen wird. Dies spricht für die Standardisierung, die Vereinheitlichung der Methoden und die Möglichkeit überregionaler klinischer Studien, da über eine solche Vorgehensweise auch geringere Effekte auf den Patientenzustand zuverlässiger erkannt werden können.
Eine wichtige Motivation für das laut Morris neue Paradigma ist die Identifikation der wirksamsten Therapien, mit dem Vorteil, dass nach der Identifikation der bestmöglichen Behandlung durch aufwendige Studien diese prinzipiell jedem Patienten durch Anwendung des Protokolls zuteil werden kann.
Betrachtet man die Standardisierungs-Bestrebungen aus Sicht der Fehlervermeidung, so ist anzumerken, dass, solange für die Behandlung von Patienten keinerlei Standardisierung existiert, es bedeutend schwerer fällt, fehlerhafte Abweichungen vom eigentlich intendierten Behandlungsplan festzustellen. Indem der Behandlungsplan in der Form expliziter Protokolle formalisiert wird, werden also weitaus bessere Möglichkeiten geschaffen, menschliche Fehler frühzeitig zu erkennen, oder Ursachen für Fehler im Nachhinein durch Analyse von elektronischen Patientenakten aufzuspüren.

1.3. Die Entwicklung medizinischer Behandlungsprotokolle

Die standardisierte Behandlung von Patienten nach zuvor auf Wirksamkeit hin validierten Protokollen, wie sie in Abschnitt 1.2 vorgestellt wird, setzt zunächst die Entwicklung solcher Behandlungsprotokolle voraus. Laut Morris [Mor98] sind auch hierfür Tests im klinischen Umfeld ein unverzichtbares Mittel. Genauer schlägt er eine iterative Verfeinerung des Modells vor, bei der das Modell bereits recht früh in der klinischen Umgebung, das heißt unmittelbar am Krankenbett, zur Generierung von Therapievorschlägen verwendet wird. Akzeptiert der Experte vor Ort die Vorschläge des Protokolls nicht, so wird die Ursache für die abweichende Meinung gesucht und gegebenenfalls das Protokoll entsprechend neuer Erkenntnisse abgeändert. Dies ist nicht mit klinischen Studien zu verwechseln, die die Wirksamkeit von Therapien oder ganzen Protokollen nachweisen (vgl. Abschnitt 1.2). Dennoch ist auch dies ein aufwendiger Ansatz. Besonders in frühen Stadien sind noch zahlreiche Fehler und großer Veränderungsbedarf am Modell zu erwarten, weshalb die Zweckmäßigkeit eines klinischen Einsatzes früher Versionen eines solchen Protokolls bezweifelt werden kann.
In dieser Diplomarbeit wird die Erstellung und Validierung von intensivmedizinischen Behandlungsprotokollen aus dem Blickwinkel der Informatik betrachtet. Die Protokolle stellen demnach eine spezielle Form wissensbasierter Systeme dar, oder lassen sich zu-

mindest, falls die bei der Entwicklung verwendete Repräsentation nicht maschinenlesbar ist, durch geeignete Übertragung in einen anderen Formalismus in ein wissensbasiertes System überführen. Solche Systeme sind seit längerem Forschungsgegenstand im Bereich der Künstlichen Intelligenz.
In erster Linie soll in dieser Arbeit eine Methode vorgestellt werden, die besonders das Finden von Fehlern und Inkonsistenzen in medizinischen Behandlungsprotokollen vereinfachen soll. Außerdem schafft sie einen geeigneten Rahmen zur Revision des Modells, um einmal gefundene Fehler auszumerzen. Bei dieser Methode wird zunächst eine von Rechnern interpretierbare Repräsentation des Behandlungsprotokolls erstellt, die dann dazu verwendet wird, das Protokoll anhand automatisch am Krankenbett erfasster Patientendaten zu validieren. Zentral ist dabei, dass sich aus den erfassten Patientendaten sowohl die jeweils aktuellen Patientenzustände, als auch die durchgeführten therapeutischen Maßnahmen ablesen lassen.

1.4. Medizinische Grundlagen

Die für diese Diplomarbeit relevante medizinische Domäne ist die Intensivmedizin. Insbesondere das hämodynamische System[1], Atmung und Lunge, sowie der Flüssigkeitshaushalt und die Nierenfunktion des Körpers sind für die nachfolgenden Kapitel von Interesse.
Es sei darauf hingewiesen, dass die Intensivmedizin in dieser Diplomarbeit hauptsächlich als Anwendungsgebiet für eine Methode der Informatik betrachtet wird. Das heißt, anhand einer exemplarischen Anwendung, die ihrerseits jedoch Praxisrelevanz besitzt, wird der Wert einer Methode evaluiert. Aus diesem Grund genügt es, wenn an dieser Stelle einige der wichtigsten verwendeten Konzepte aus der Medizin kurz erläutert werden, um einen Eindruck des für die Modellierung relevanten Sachbereichs und dessen Komplexität zu vermitteln.
Eine Übersicht über verschiedene Ziele einer intensivmedizinischen Betreuung, sowie über dabei zum Einsatz kommende Methoden gibt Tabelle 1.1. Sie bezieht sich speziell auf die Stabilisierung chirurgischer Hochrisikopatienten nach Operationen. Von den medizinischen Begriffen können hier nur einige aufgegriffen werden. Eine umfassendere Übersicht geben beispielsweise [AE97] oder [Hil98], eine ausführliche Beschreibung von Medikamentenwirkungen findet sich in [Mut97].

1.4.1. Herz und Kreislaufsystem

Die Hauptaufgabe des Herz-/Kreislaufsystems ist die Versorgung des Körpers mit Sauerstoff und anderen Nährstoffen. Bei der Versorgung von Intensivpatienten ist es von zentraler Bedeutung, den Kreislauf stabil zu halten. Zur Bewertung der Effektivität des Kreislaufs gibt es zahlreiche Parameter, von denen einige im folgenden vorgestellt werden.
Das Herz fungiert als Pumpe des Körpers. Durch den Herzschlag zirkuliert das Blut, was die Versorgung des Körpers mit Sauerstoff und Nährstoffen überhaupt erst ermöglicht. Die Herzfrequenz gibt dabei die Anzahl der Herzschläge an, üblicherweise in der Einheit „Schläge pro Minute“. Eine zu hohe Herzfrequenz wird als „Tachykardie“ bezeichnet.
Mit dem Herzzeitvolumen ist die durch das Herz gepumpte, also die vom Herzen ausgeworfene Menge Blut gemeint. Diese wird üblicherweise in der Einheit „Liter pro Minute“

[1] Das hämodynamische System umfasst den Blutkreislauf mit seinen physikalischen Strömungseigenschaften.

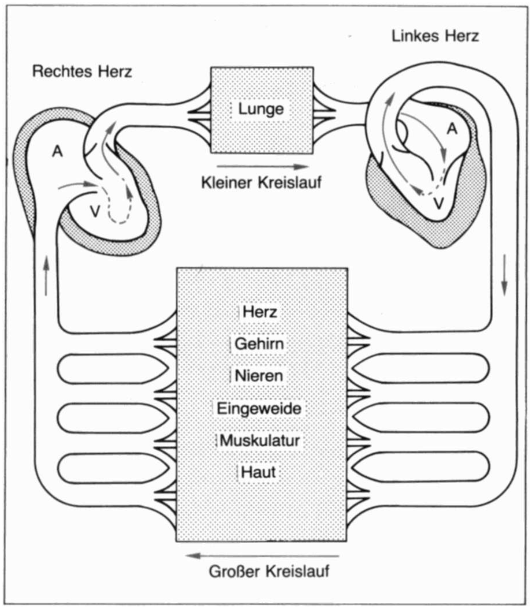

Abbildung 1.1.: Körper- und Lungenkreislauf. Quelle: [TMV91]

notiert. Um eine von Gewicht und Größe des Patienten weitestgehend unabhängige Normierung dieses Wertes zu erhalten, kann statt dem Herzzeitvolumen auch der Quotient aus Herzzeitvolumen und Körperoberfläche herangezogen werden. Der resultierende Parameter wird als *Herzindex* (CI) bezeichnet.

Der Kreislauf des Körpers besteht aus zwei getrennten Teilkreisläufen, der Lungenstrombahn (kleiner Kreislauf) und dem Körperkreislauf (großer Kreislauf). Dementsprechend ist das Herz, wie in Abbildung 1.1 dargestellt, zweigeteilt. Vom rechten Herzen aus wird das sauerstoffarme Blut durch die Lungenstrombahn geführt. Beim Passieren der Lunge wird es mit Sauerstoff angereichert. Vom linken Herzen aus strömt das sauerstoffreiche Blut unter hohem Druck über die Aorta und die herznahen Arterien in den Körperkreislauf. Im weiteren Verlauf verästeln sich die Gefäße bis zu den Kapillaren, die die Organe durchbluten und mit Nährstoffen und Sauerstoff versorgen. Das im Anschluss sauerstoff- und nährstoffarme Blut strömt weiter in wieder zunehmend größere Gefäße, bis es schließlich unter sehr niedrigem Druck über große Venen das rechte Herz erreicht. Sowohl das linke als auch das rechte Herz unterteilen sich weiter in einen Vorhof (Atrium) und eine Herzkammer (Ventrikel), wobei das Blut vom Vorhof in die Herzkammer, und von dort in den Kreislauf gelangt.

Der arterielle Blutdruck bezeichnet den Druck in den herznahen Arterien des Körperkreislaufs. Dieser Druck schwankt zyklisch, je nachdem, ob sich das Herz gerade mit Blut füllt, oder ob eine Kontraktion des Herzens stattfindet, die das Blut von der Herzkammer in den Kreislauf pumpt. Aus diesem Grund unterscheidet man zwischen dem höheren *systolischen* und dem niedrigeren *diastolischen* arteriellen Blutdruck. Ein Maß, das die beiden Formen des Blutdrucks vereint, ist der *mittlere arterielle Blutdruck*. Dieser Wert ergibt sich als der durchschnittliche Blutdruck über einen Herzzyklus, gewichtet also den systolischen und diastolischen Blutdruck entsprechend. Ist der Druck zu niedrig, so spricht man von „Hypotonie", ein zu hoher Druck wird als „Hypertonie" bezeichnet. Eine angenehme Eigenschaft des mittleren arteriellen Drucks ist, dass dieser robuster gegen

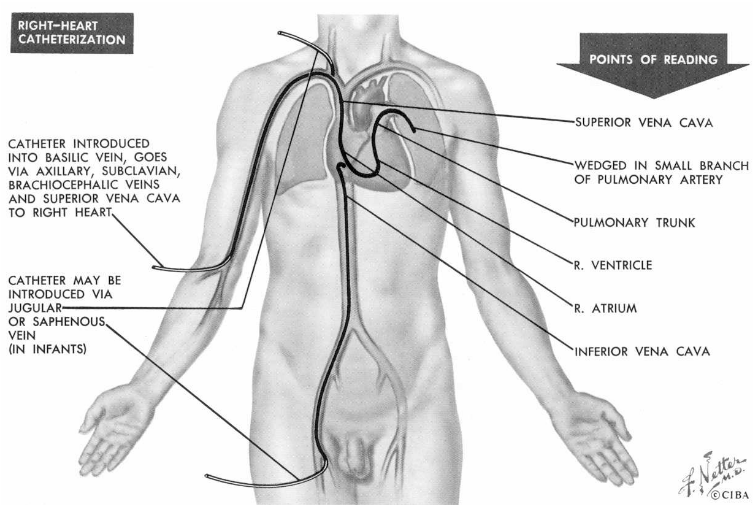

Abbildung 1.2.: Rechtsherz-Katheterisierung. Quelle: [Net69]

die Art der Messung ist, als der systolische oder diastolische Druck, das heißt, Werte die über eine invasive Messung (Gefäßkatheter) ermittelt werden unterscheiden sich nicht wesentlich von über eine Blutdruckmanschette gemessenen Werten.
Zu den im Brustkorb gemessenen Blutdrücken zählen der zentralvenöse Druck, und der pulmonalkapilläre Verschlussdruck. Beide Drücke werden im venösen Teil des Kreislaufs gemessen und sind deshalb deutlich niedriger als die Drücke im arteriellen Teil des Kreislaufs.
Der *zentralvenöse Blutdruck* (ZVD) bezeichnet den Blutdruck in den großen Venen unmittelbar vor dem rechten Vorhof. Diesem Druck sehr ähnlich ist der Druck im Vorhof des rechten Herzens, der sogenannte *rechtsatrielle Blutdruck*.
Als *pulmonalkapillären Verschlussdruck* (PCWP) bezeichnet man den Druck, der gemessen werden kann, wenn ein kleiner Ast der Lungenarterien mittels eines Katheters verschlossen wird und so eine stehende Blutsäule bis zum linken Vorhof geschaffen wird. Er wird als ein Maß des Drucks im linken Vorhof angesehen. Dieser Druck kann nur sporadisch durch manuelle Messung mithilfe eines Rechtsherzkatheters ermittelt werden. Aus diesem Grund ist dieser Messwert nur selten verfügbar. Abbildung 1.2 skizziert das Vorgehen bei einer Rechtsherz-Katheterisierung.
Um einen stabilen Kreislauf zu ermöglichen, muss sich eine geeignete Menge Volumen im Kreislauf befinden. Eine wesentliche, den Kreislauf beeinflussende Größe, die vom Körper auch beispielsweise zur Regulation des Blutdrucks verwendet wird, ist der *totale periphere Widerstand*, also der Strömungswiderstand im Körperkreislauf. Dieser resultiert aus den Strömungswiderständen der vielen einzelnen Gefäße. Der mittlere arterielle Blutdruck, eine der wesentlichen Größen in der medizinischen Versorgung, ergibt sich unmittelbar aus dem Herzzeitvolumen und dem totalen peripheren Widerstand. Diesem Zusammenhang liegen einfache physikalischen Gesetzmäßigkeiten zugrunde.
Zur Regulation der Durchblutung kann der Körper verschiedene Gefäße eng oder weit stellen, und auf diese Weise nicht nur den arteriellen Mitteldruck ändern, sondern darüber hinaus auch gezielter die Versorgung der verschiedenen Organe steuern. Die Fließeigenschaften des Blutes hängen jedoch auch von bedeutend komplexeren Faktoren ab, etwa von der Zähflüssigkeit des Blutes (Viskosität), oder davon, ob sich im Blutstrom Verwir-

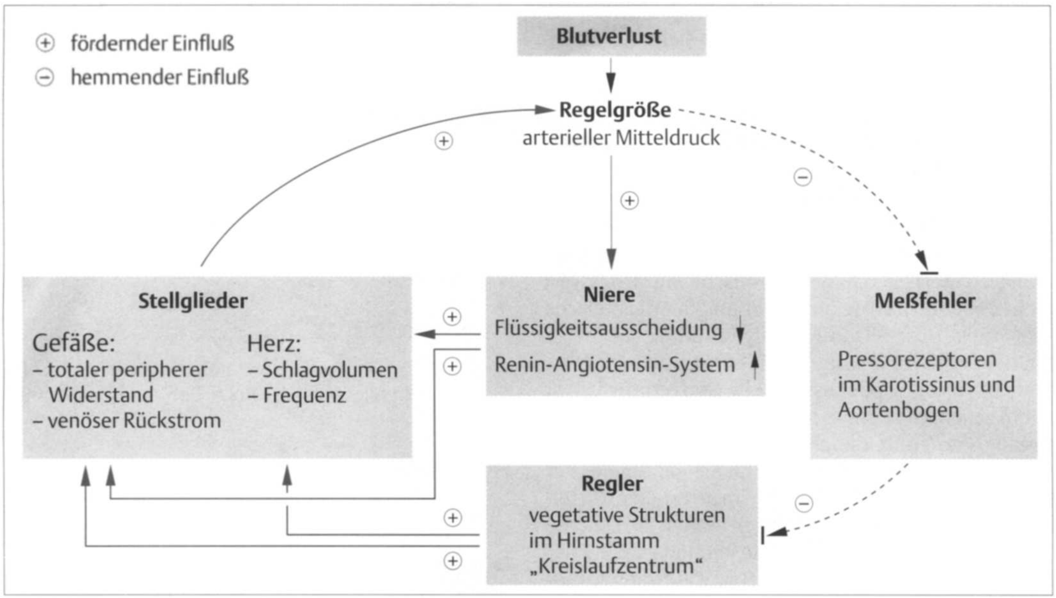

Abbildung 1.3.: Regelkreis für die kurzfristige Regulation des Blutdrucks. Quelle: [AE97]

belungen ergeben.
Der Körper besitzt verschiedene Mechanismen, um den Kreislauf zu regulieren. Einer davon baut auf Rezeptoren auf, die direkt auf den arteriellen Blutdruck und dessen Schwankungen und Änderungen in den großen Gefäßen des Körperkreislaufs reagieren (Pressorezeptorenregelkreis). Hierbei wird jeweils über das Kreislaufzentrum im Hirnstamm der Änderung entgegengewirkt, etwa durch Weitstellung von Gefäßen (Vasodilation) oder der Verringerung der Herzfrequenz. Ein weiterer Mechanismus reagiert auf Minderdurchblutung des Gehirns, indem über das Kreislaufzentrum der mittlere arterielle Blutdruck erhöht wird (Ischämie-Reaktion des Zentralen Nervensystems). Schließlich gibt es noch mehrere längerfristige Regelungsmechanismen für den Kreislauf, beispielsweise eine Änderungen des im Kreislauf befindlichen Gesamtvolumens über die Niere. Dies kann entweder durch vermehrte Ausscheidung, oder aber durch Zurückhaltung und Umverteilung von Flüssigkeit geschehen. Eine Übersicht der wichtigsten Regulationsmechanismen findet sich in Abbildung 1.3.

1.4.2. Atmung

Die Lungenstrombahn ist der Teil des menschlichen Kreislaufs zwischen rechtem und linkem Herzen. Sie dient der Anreicherung des Blutes mit Sauerstoff, das dann über den Körperkreislauf die Organe versorgt, und der gleichzeitigen Abatmung von Kohlendioxid, als einem Stoffwechselendprodukt des Körpers. Der Gasaustausch in der Lunge basiert dabei auf dem Prinzip der Konzentrationsangleichung der Gase im Blut und der Atemluft. Im Normalfall wird die Atmung, gemäß des CO_2-Anteils im Blut, über das Atemzentrum des Gehirns gesteuert. Bei unzureichender eigener Atmung müssen Beatmungsgeräte eingesetzt werden, da dies sonst nach kürzester Zeit zur Schädigung von Organen führt. Bei einer künstlichen Beatmung mit entsprechenden Geräten kann, in Abhängigkeit vom Krankheitsbild, zwischen verschiedenen Atemfrequenzen, Beatmungsmustern und Beatmungsdrücken gewählt werden, etwa um den Gasaustausch zwischen Blut und Lunge zu verbessern. Da dieser Bereich in den folgenden Kapiteln nur am Rande berührt wird, soll hier auf eine ausführlichere Beschreibung verzichtet werden.

1.4.3. Wasser- und Elektrolythaushalt

Der menschliche Körper ist sehr bestrebt, sowohl die Flüssigkeitsmenge, als auch die Elektrolyt-Konzentrationen in bestimmten Grenzen zu halten (Homöostase) und reagiert sehr empfindlich auf Störungen des Wasser-/Elektrolythaushalts. Für diese Arbeit sind besonders der Wasser- und der Natriumhaushalt interessant.
Regelungsmechanismen für die Aufrechterhaltung des Wasserhaushalts sind besonders an die Tätigkeit der Niere geknüpft. Je nach Steuerung der Nierentätigkeit über ein antidiuretisches Hormon kann entweder Wasser verstärkt ausgeschieden, oder im Körper belassen werden. Die Steuerung des Haushalts geschieht über spezielle Rezeptoren, die auf die Teilchenkonzentration des Blutes (Osmolalität) reagieren.
Der Natriumhaushalt wird ebenfalls in besonderem Maße über die Niere reguliert. Der Flüssigkeitshaushalt hängt vom Natriumhaushalt dadurch ab, dass das Blutplasmavolumen[2] sich in Abhängigkeit von der Natriumkonzentration ändert. Die Aufnahme von Natrium erfolgt hauptsächlich über die Nahrung. Die Ausscheidung über die Niere erfolgt derart, dass eine Normalisierung sowohl des Flüssigkeits-, als auch des Natriumhaushalts eintritt.

1.4.4. Interventionen

In diesem Abschnitt sollen einige wichtige Interventionen vorgestellt werden, die in der Intensivmedizin bei der Stabilisierung von Patienten Anwendung finden. Eine ausführliche Beschreibung vieler Medikamente findet sich beispielsweise in [Mut97].
Vasokonstringenzien sind Medikamente mit gefäßverengender Wirkung. Die Engstellung der Gefäße nach der Gabe eines Medikamentes dieser Art hat eine Steigerung des arteriellen Mitteldrucks zur Folge.
Unter *Vasodilatoren* fallen Medikamente, die gefäßerweiternd wirken. Die Gabe eines Medikamentes dieser Art senkt den arteriellen Mitteldruck.
Positiv inotrope Medikamente sind solche, die die Herzkraft steigern, also die Schlagstärke oder Kontraktionskraft des Herzens erhöhen.
Diuretika sorgen für eine verstärkte Ausscheidung von Flüssigkeit über die Niere (forcierte Diurese). Auf diese Wiese wird dem Körper Flüssigkeit entzogen, was langfristig auch Auswirkungen auf den Kreislauf hat.
Schließlich können Infusionen, und zwar dann wenn sie schnell laufen, eine kreislaufstabilisierende Wirkung haben, da dem Kreislauf in diesem Fall in kurzer Zeit Volumen zugeführt wird.

1.5. Verwandte Arbeiten

1.5.1. Arbeiten des Sonderforschungsbereiches 475, Teilprojekt C4

Ein besonderes Problem des verantwortlichen Personals moderner Intensivstationen ist die große Menge erfasster Patientendaten. Während die zunehmende Genauigkeit und

[2] Blutplasma bezeichnet den flüssigen Anteil des Blutes

Organ	Intervention	Ziele	Messwerte
Herz **Kreis-lauf**	• positiv inotrope Substanzen (Erhöhung der Herzleistung) • Erhöhung der Vorlast (Erhöhung der Herzfüllung) • Senkung der Vorlast (Verminderung der Herzfüllung) • Erhöhung der Nachlast (Erhöhung der Gefäßwiderstände) • Verminderung der Nachlast (Verminderung der Gefäßwiderstände)	• Herzzeitvolumen > 4.5 • Arterieller Mitteldruck 90 - 100 mmHG • Herzfrequenz 60 - 100/min	• Herzzeitvolumen • Herzfrequenz • Arterieller Blutdruck • Pulmonal.-art. Blutdruck • zentralvenöser Blutdruck • PCWP • Gefäßwiderstände • Schlagleistung
Lunge	• Sauerstoffgabe • Beatmung • Recruitment von Alveolen durch PEEP • erhöhte Beatmungsdrücke • Inverse Beatmung • Bauchlagerung	• $p_aO_2/F_iO_2 > 300$ mmHg • Beatmungsdrücke niedrig • Atemzeitverhältnis $< 1 : 2$	• p_aO_2, p_aCO_2, pH • Beatmungsdrücke • Atemzugvolumen • F_iO_2 • Atemzeitverhältnis
Niere	• Volumengabe (Erhöhung des Flüssigkeitsangebotes) • Erhöhung des Blutdrucks (Erhöhung des renalen Perfusionsdruckes) • Erhöhung des HZV (Erhöhung der renalen Perfusion) • Dopamingabe mit 1 - 3 μg/kg/min (Erhöhung der renalen Perfusion)	• Urinproduktion > 100 ml/h (primäres Ziel) • Creatinin i.S. < 1.0 • Harnstoff i.S. < 50	• Urinproduktion • Creatinin i.S. • Harnstoff i.S.

Tabelle 1.1.: Therapieziele bei der Behandlung chirurgischer Hochrisikopatienten. Quelle: [The99]

die große Menge verschiedener überwachter Parameter dazu führen, dass der Patientenzustand immer besser abgebildet wird, ergibt sich nun ein neues Problem für den menschlichen Entscheidungsträger. Bei einer großen Fülle oft kontinuierlich (etwa minütlich) erfasster Daten fällt es schwerer, relevante Informationen zu erkennen. Bei vielen relevanten Faktoren, die gleichzeitig in einem Entscheidungsprozess zu berücksichtigen sind, stoßen Menschen schnell an ihre natürlichen Grenzen. So ist es Menschen nicht möglich, mehr als 7 (maximal 9) unabhängige, relevante Variablen gleichzeitig in einem Entscheidungsfindungsprozess zu berücksichtigen [Mil56].
Da es mit zunehmendem Umfang der Datenerfassung für menschliche Entscheidungsträger immer schwieriger wird, alle relevanten Aspekte im Auge zu behalten, ist es ein interessanter Forschungsbereich, die Komplexität der Daten zu reduzieren. Mit dieser Aufgabe beschäftigt sich das Teilprojekt C4 des Sonderforschungsbereichs 475. Die Daten haben dabei die Form multivariater Zeitreihen. In diesem Rahmen ist ein hybrider Ansatz entstanden, der sowohl mit Methoden des Maschinellen Lernens aus den protokollierten Verlaufsdaten ein Modell des Interventionsverhaltens von Ärzten herzuleiten versucht, als auch über deklarative Regeln modelliertes medizinisches Wissen verwendet [MIB+00]. Auf diese Weise kann medizinisches Hintergrundwissen dazu verwendet werden, zu validieren, ob die vom gelernten Modell vorgeschlagenen Interventionen tatsächlich den Patienten stabilisieren würden.
Statistische Methoden helfen bei diesem Ansatz nicht nur, bei der Messung aufgetretene Fehler zu erkennen ([BGI99]), also die Datenqualität zu verbessern, sondern darüber hinaus werden die Daten in eine abstraktere Form überführt. Diese Form besteht aus zeitlichen Intervallen, die den Verlauf von Parametern qualitativ wiedergeben, wie etwa „der mittlere arterielle Blutdruck stieg von Minute 12 bis Minute 39 an“. Auf diese Weise wird eine Brücke zwischen der Zeitreihenanalyse und symbolischen Verfahren geschlagen (siehe auch [MBJ99]). In [IBGL97] findet sich eine Untersuchung zur Qualität dieses Abstraktionsschrittes im Umfeld intensivmedizinischer Zeitreihen.
Ein Nachteil der in [MIB+00] automatisch akquirierten Modelle zum Interventionsverhalten besteht darin, dass diese Modelle die Form direkter Abbildungen von Messwerten auf Interventionen haben, die wiederum nur im Herauf- oder Herabsetzen von Medikamentendosierungen bestehen („State-Action“-Regeln). Ärzte gehen bei der Behandlung eher planbasiert vor, das heißt Entscheidungen werden immer in einem Kontext, beispielsweise im Kontext der Krankengeschichte oder voriger Interventionen gesehen. Therapien werden längerfristig angesetzt und sollen dazu dienen, dass sich der Patientenzustand kontinuierlich vorteilhaft entwickelt. Bei „State-Action“ hingegen bezieht sich die vorgeschlagene Intervention ausschließlich auf den letzten erfassten Zeitabschnitt, es werden also weder zeitliche Entwicklungen bewertet, noch gibt es einen längerfristigen Behandlungsplan.
Das in dieser Diplomarbeit validierte Behandlungsprotokoll basiert auf für Menschen verständlichen Regeln, berücksichtigt zeitliche Verläufe und komplexere Begriffe[3]. Daher spiegelt dieses Protokoll eher ärztliches Vorgehen wider, als „State-Action“-Regeln.

[3] „Komplexere Begriffe“ meint hier Begriffe, deren Definition sich aus mehreren, direkt gemessenen Basisbegriffen zusammensetzt.

1.5.2. Verarbeitung intensivmedizinischer Zeitreihen

Neben den in Abschnitt 1.5.1 erwähnten Ansätzen werden noch weitere Methoden zur Vorverarbeitung speziell intensivmedizinischer Zeitreihen diskutiert, entweder zur Identifikation von Messfehlern, oder zur Ableitung symbolischer Repräsentationen.
In [TKM00] wird eine Methode vorgestellt, mit der der hohe Anteil automatisch von Messgeräten ausgelöster Fehlalarme in der Intensivstation[4] verringert werden kann. Alarme, die durch fehlerhafte Messungen hervorgerufen werden, werden von einem zuvor in einer Trainingsphase aus klassifizierten Beispielen gelernten Entscheidungsbaum experimentell recht zuverlässig erkannt. Dabei wird von der Tatsache Gebrauch gemacht, dass einige der gemessenen Variablen nicht unabhängig voneinander sind. Um Messfehler einer Variablen zu identifizieren, beziehen die Bäume erfolgreich andere Variablen mit ein.
In [CCPR00] wird eine von der in Abschnitt 1.5.1 verschiedene Methode zur Erkennung von Trends in einer intensivmedizinischen Zeitreihe vorgestellt. Hierbei liegt besonderes Gewicht auf der Erkennung von Ausreißern bei der Messung und Robustheit gegenüber Rauschen.
In [KMSK92] geht es um die Gewinnung einer symbolischen Repräsentation aus intensivmedizinischen Zeitreihen, speziell zur Konstruktion eines hämodynamischen Monitoringsystems. Die Beseitigung von Ausreißern wird mit einem Median-Filter durchgeführt und die quadratischen Abweichung von prognostizierten Trends dienen als Maß für deren Zuverlässigkeit. Die erhobenen Werte und die Trends werden anhand fester Intervalle diskretisiert. Die Intervalle stellen medizinisches Wissen über die verarbeiteten klinischen Variablen dar. Die Qualität dieses Verfahrens zur Erkennung von Trends wurde mit Hilfe eines Mediziners abgeschätzt.
Die Ableitung intervallbasierter Abstraktionen aus medizinischen Zeitreihen ist auch die in [SM96] behandelte Aufgabe. Es werden dazu wiederkehrende Teilaufgaben der Abstraktion identifiziert, für die allgemein gehaltene, wissensbasierte Methoden vorgestellt werden. Diese Methoden können durch die Ergänzung spezifischen Domänen-Wissens umfangreich eingesetzt werden. Verschiedene Ontologien dienen der Organisation des für die Abstraktion benötigten Wissens.
Ein ganz anderer Ansatz wird in [SLDE97] beschrieben. Es wird versucht, im intensivmedizinischen Umfeld typische hämodynamische Muster, also zeitlich aufeinander folgende Ereignisse im hämodynamischen System, mit Hilfe Neuronaler Netze zu erkennen, und schließlich sogar vor der vollständigen Beendigung Sequenzen vorhersagen zu können.

[4] Der Artikel bezieht sich speziell auf die Neugeborenen-Intensivstation.

2. Eine Entwicklungsumgebung für Behandlungsprotokolle

Wie im letzten Kapitel dargestellt wurde, bieten rechnerbasierte Behandlungsprotokolle neue Chancen für die Medizin. Die Entwicklung der benötigten Protokolle ist jedoch aufwendig. So nahm die Entwicklung des in Abschnitt 1.2 bereits erwähnten Beatmungsprotokolls für das Krankheitsbild des akuten Lungenversagens 25 Personenjahre in Anspruch. Wie in Abschnitt 1.3 erwähnt, wurde dieses Protokoll bereits früh im klinischen Betrieb eingesetzt und aufgrund der dort gemachten Erfahrungen laufend verbessert. Eventuell stellt die frühe, recht starke Bindung an den klinischen Einsatz einen der Hauptgründe für die lange Entwicklungszeit dar. Hier könnte ein alternativer Ansatz helfen, der stärker auf informationsverarbeitende Technologien setzt.

2.1. Vorteile maschinenlesbarer Repräsentationen

Während die Entwicklung medizinischer Behandlungsprotokolle im klinischen Umfeld prinzipiell auch mit einer menschen- aber nicht maschinenlesbaren Repräsentation des Protokolls auskommt, ist für die Nutzung vieler rechnerbasierter Werkzeuge die Maschinenlesbarkeit unabdingbar. Solche Werkzeuge können die Modellierung des für die Aufgabe relevanten Wissens erleichtern. Erst wenn ein Rechner das modellierte Wissen interpretieren kann, lassen sich automatisch Konsequenzen des Modells aufzeigen. Insbesondere Widersprüche können so leichter entdeckt werden. Nützlich ist auch eine Unterstützung bei Korrekturen. Begreift man die Modellierung eines Sachbereichs als zyklischen Prozess, bei dem in jedem Fall mit Revisionen zu rechnen ist (vgl. [Mor89]), so kommt einer Umgebung mit geeigneten Werkzeugen zur Wartung und Revision eine besondere Bedeutung zu.

Ein Behandlungsprotokoll als maschinenlesbare Wissensbasis zu repräsentieren beinhaltet eine Explizierung des beabsichtigten Protokollverhaltens in einem geeigneten Formalismus. Eine solche Darstellung erhöht durch Vermeidung jeglicher Ungenauigkeiten die Diskutierbarkeit des Protokolls (siehe auch Abschnitt 1.2). Sollen verschiedene Protokolle miteinander verglichen werden, die in geeigneter Weise formalisiert wurden, so können für die Modelle sogar sinnvoll formale Ähnlichkeitsmaße definiert werden. Ein solcher Ansatz wird in [Müh94] dazu verwendet, die „Distanz" zwischen verschiedenen kindlichen Erklärungen für den Tag/Nachtzyklus zu untersuchen. Nach einer Übersetzung der Erklärungen in formale logische Modelle gelang die Anwendung von Ähnlichkeitsmaßen, die auf den Operatoren „Fakt hinzufügen", „Fakt löschen" und „Wissensrevision" basieren.

2.2. Die Verwendung protokollierter Verlaufsdaten

Um die Qualität eines Behandlungsprotokolls zu evaluieren, kann dieses testweise im klinischen Umfeld eingesetzt werden. Vom Aufwand her günstiger erscheint jedoch, zumindest in frühen Stadien erst einmal ganz auf klinische Tests zu verzichten, und statt dessen die Angemessenheit anhand *protokollierter* Patientendaten zu validieren. In der modernen Intensivmedizin ist eine elektronische Aufzeichnung der erhobenen Verlaufsdaten oft ohnehin gegeben, so dass entsprechende Daten in ausreichender Menge zur Validierung verfügbar sind.

Ob die Methode geeignet ist, um intensivmedizinische Wissensbasen zu validieren, oder gar um diese (weiter) zu entwickeln ist Gegenstand dieser Diplomarbeit. Sind neben den für die Entscheidung relevanten Verlaufsdaten auch die durchgeführten medizinischen Interventionen in elektronischer Form verfügbar, so könnte das in Abschnitt 1.3 beschriebene Modell der sukzessiven Verbesserung im klinischen Umfeld wie folgt am Rechner nachgeahmt werden:

Die Wissensbasis wird dazu verwendet, unabhängig von den im Datensatz vermerkten Interventionen eine Therapie vorzuschlagen. Weichen Vorschläge und tatsächlich getätigte Interventionen voneinander ab, so werden die entsprechenden Fälle näher untersucht. Lässt dies Unzulänglichkeiten der Wissensbasis erkennen, so wird ein Revisionsschritt durchgeführt.

Die Suche nach Widersprüchen zwischen einem modellierten Behandlungsprotokoll und den protokollierten klinischen Entscheidungen lässt sich durch die Angabe einiger weiterer Regeln automatisieren, was einen deutlichen Vorteil gegenüber der Verbesserung des Protokolls im klinischen Umfeld darstellt. Ist der verfügbare Datenbestand groß genug, so kann in kurzer Zeit die Angemessenheit des Protokolls automatisch evaluiert werden. Entsprechend können die Auswirkungen von Revisionen des Modells abgeschätzt werden. Zumindest während früher Entwicklungsstadien eines Protokolls für eine komplexe Domäne dürften sich mit der beschriebenen Methode noch leicht Fehler und Unzulänglichkeiten finden lassen. In diesem Stadium ist zu erwarten, dass der Aufwand eines solchen Vorgehens bedeutend geringer ist, als wenn alle Fehler im Protokoll erst im klinischen Einsatz zu Tage träten und womöglich ohne die Unterstützung durch geeignete Werkzeuge korrigiert werden müssten.

Auch hinsichtlich der Weiterentwicklung eines Protokolls ist das beschriebene Vorgehen vielversprechend. Experten wird ein Werkzeug an die Hand gegeben, mit dem sich verschiedene Alternativen für ein Behandlungsprotokoll mit geringem Aufwand „durchspielen“ lassen. Konsequenzen von Änderungen am Modell können mit diesem Hilfsmittel direkt über den Verlaufsdaten einer Vielzahl von Patienten abgeschätzt werden. Die Verwendung realer Patientendaten lässt dabei hoffen, oft ähnlich gute Ergebnisse wie bei bedeutend aufwendigeren klinischen Tests zu erhalten.

Darüber hinaus lässt ein einmal geeignet operationalisiertes Protokoll Experimente zu, die für die Entwicklung entsprechender Wissensbasen von grundsätzlichem Interesse sind. Ein in verwandten Arbeiten immer wieder diskutiertes Thema ist die Vorverarbeitung von Zeitreihen. Eine Zielsetzung dabei ist die Verbesserung der Datenqualität, beispielsweise indem bei der Messung aufgetretene Artefakte[1] herausgefiltert werden. Mit Hilfe eines operationalen Protokolls kann experimentell untersucht werden, wie sich das Verhalten (besonders die Interventionsvorschläge) nach unterschiedlichen Vorverarbeitungsschritten ändert.

[1]Gemeint sind besonders kurzzeitige Messfehler.

3. Repräsentation des Behandlungsprotokolls

3.1. Bemerkungen zum Vorgehen

Die im letzten Kapitel vorgestellte Methode basiert auf einem operationalen Behandlungsprotokoll. Da die Entwicklung eines solchen Protokolls, ohne auf Vorarbeiten zurückzugreifen, im Rahmen einer Diplomarbeit nicht möglich ist, wird in dieser Arbeit ein bestehendes, noch nicht operationales Protokoll verwendet.

Das Thema dieses Kapitels ist die Operationalisierung des Protokolls durch die Repräsentation in einem geeigneten Formalismus. Dabei ist eine deklarative Form der Repräsentation wichtig. Medizinische Experten sollten ohne größere Schwierigkeiten imstande sein, die Herleitungen von Interventionsvorschlägen nachvollziehen zu können. Die resultierende Wissensbasis soll mit geringem Aufwand wartbar sein und allgemeinen Anforderungen an Wissensbasen, wie Übersichtlichkeit des repräsentierten Wissens und Modularität gerecht werden.

Darüber hinausgehende Anforderung an die Operationalisierung ergeben sich aus dem Ziel, einen Rahmen für Experimente zu schaffen. Dies äußert sich in Effizienz-Anforderungen an das resultierende System. Die Interpretation des repräsentierten Wissens sollte sich in einem zeitlichen Rahmen abspielen, der es erlaubt, mit einem Experten vor Ort verschiedene Alternativen für ein Modell durchzuspielen. Weiterhin zentral für Experimente ist der Zugriff auf Verlaufsdaten. Dieser sollte vom laufenden System aus möglich sein.

Im folgenden Abschnitt wird zunächst das verwendete Protokoll beschrieben. Im Anschluss daran wird allgemein auf die Erstellung wissensbasierter Systeme eingegangen und darauf aufbauend die gewählte Art der Repräsentation motiviert. Den Schwerpunkt bildet eine Beschreibung der entstandenen Wissensbasis.

3.2. Das verwendete Protokoll

Für diese Diplomarbeit stand das unfertige Protokoll einer Forschergruppe aus den USA zur Verfügung. Es handelt sich dabei um ein intensivmedizinisches Behandlungsprotokoll, das sich auf hämodynamische Aspekte einschließlich des Flüssigkeitshaushalts beschränkt.

Da das Protokoll unveröffentlicht ist und um Missverständnissen bezüglich der in dieser Arbeit erzielten Ergebnisse vorzubeugen, wird hier bewusst auf die Angabe einer Quelle verzichtet. Weiterhin wird auf die nicht zweckdienliche Veröffentlichung von Details in dieser Arbeit verzichtet.

3.2.1. Eigenschaften des Protokolls

Das Protokoll liegt in einer Papierversion vor. Es handelt sich dabei um ein Flussdiagramm mit den üblichen Elementen, also Abfragen, die gleichzeitig Verzweigungen darstellen, und Anweisungen. Zu den Anweisungen zählen Variablenzuweisungen oder Anweisungen an das medizinische Personal, etwa eine Messung vorzunehmen oder eine therapeutische Maßnahme durchzuführen. Innerhalb des Flussdiagramms wird Gebrauch von einem aussagenlogischen Formalismus gemacht. Da das Modell in der vorliegenden Form selbstverständlich nicht maschinenlesbar ist, ist eine Übertragung des Modells in eine andere Repräsentation unverzichtbar.
Das Modell gibt die Sicht eines behandelnden Arztes wieder. Es wird also nicht minütlich die Medikation angepasst, sondern die Behandlung wird durch regelmäßige Visiten jeweils längerfristig festgesetzt. Jede Visite entspricht einem Durchlauf durch das Protokoll. Je nach Verfassung des Patienten liegt die maximal zulässige Zeit zwischen zwei Visiten zwischen 30 Minuten und 4 Stunden. Zu jeder Visite werden die neusten verfügbaren Werte der relevanten Variablen und deren Veränderung innerhalb der letzten viertel Stunde betrachtet. Darüber hinaus gibt es noch einige Ereignisse, die längerfristig von Bedeutung sind, und daher den Patientenzustand über mehrere Visiten hinweg kennzeichnen. Hierzu zählt beispielsweise das Auftreten eines sehr niedrigen Blutdrucks. Dieses Ereignis hat bis zu einen Tag lang Einfluss auf Therapieentscheidungen.
Zu den auffälligen Eigenschaften des Modells gehört, dass das knapp 50 Seiten umfassende Flussdiagramm stellenweise recht unstrukturiert wirkt und dadurch schlecht überschaubar ist. Definitionen und Anweisungen sind jedoch sehr präzise[1]. Wird ein Begriff wie „zu niedriger Blutdruck" benötigt, so wird dieser in Bezug auf eine bestimmte Variable und unter Verwendung fester Grenzwerte definiert, also beispielsweise „arterieller Mitteldruck ist kleiner als 70 mmHg". Wenn mehrere erhobene Variablen als Indikatoren für einen bestimmten Sachverhalt fungieren können, diese aber nicht unbedingt alle ständig verfügbar sind, so gibt das Protokoll eine eindeutige Hierarchie an, welcher Wert wann zu verwenden ist. Auch die Interventionsanweisungen sind sehr genau. Für Medikamente wird jeweils die genaue Dosis errechnet. Für Infusionen werden sowohl die Bestandteile vorgegeben, als auch genaue Angaben gemacht, welches Volumen in welcher Zeit durchlaufen soll.
Die vorliegende Version ist in verschiedener Hinsicht unfertig. Zunächst fällt auf, dass einige Teilaspekte vorbereitet, aber nicht ausgearbeitet sind. So ist die gezielte Erkennung und Rücknahme schädlicher Therapien offenbar erst Bestandteil späterer Versionen. Weiterhin sind einige Aspekte protokollintern widersprüchlich. Als Beispiel soll hier die teilweise Integration eines Medikamentes dienen, dessen Dosis zwar an einer Stelle hochgesetzt, die aber nirgends wieder herabgesetzt wird. Schließlich ist das Protokoll an einigen Stellen noch unpräzise, also nur nach einer das Protokoll ergänzenden Explizierung in eine formale Repräsentation übertragbar. Diese Probleme lassen sich zwar alle auf den Entwurfsstatus zurückführen, erforderten jedoch einige zusätzliche Vorarbeiten. Um eine sinnvolle Validierung zu ermöglichen, wurden mit Hilfe eines Experten für den Bereich der Intensivmedizin stimmige Anpassungen gefunden. Notwendig war zunächst, nach der Erläuterung einiger medizinischer Fachbegriffe, passende Ergänzungen für die lückenhaften Teile des Protokolls zu finden. So war etwa eine Tabelle mit der maximalen Gültigkeitsdauer für die verschiedenen erhobenen Werte noch nicht ausgefüllt.

[1] Wahrscheinlich aufgrund des Entwurfsstadiums der vorliegenden Version, fallen einige Stellen hinsichtlich Genauigkeit und Verständlichkeit aus dem Rahmen.

Im Laufe der Modellierung stellten sich einige Anweisungen des Protokolls als zu unpräzise heraus, um unmittelbar formal repräsentiert werden zu können, und mussten zunächst durch zusätzliches Expertenwissen ergänzt werden. Analog wurden Inkonsistenzen im Protokollentwurf aufgelöst. Es gelang schließlich, die vorliegende Version mit nur wenigen Änderungen für den geplanten Einsatz aufzubereiten.

3.2.2. Schnittstellen des Protokolls

Dieser Abschnitt beschreibt die Schnittstellen des verwendeten Protokolls, die sich in Eingaben und Interventionen unterteilen lassen. Die Eingaben des Protokolls sind Patientenvariablen, die innerhalb des Protokolls zur Bewertung des Zustands und zur Ableitung von Interventionen verwendet werden. Für einige dieser Variablen wird bei Bedarf eine Anweisung an das medizinische Personal erteilt, diese zu erheben. Von diesen Anweisungen abgesehen beschränkt das Verhalten des Protokolls „nach außen hin“ sich auf die Gabe von Medikamenten und Infusionen. Erläuterungen zu den medizinischen Zusammenhängen finden sich in Abschnitt 1.4.

Verwendete medizinische Variablen

Das Protokoll verwendet als Parameter für die Bewertung des Kreislaufs in erster Linie den *mittleren arteriellen Blutdruck*, den *Herzindex* und entweder[2] den *rechtsatrialen Blutdruck* oder den *pulmonalkapillären Verschlussdruck*. Außerdem wird die *Herzfrequenz* berücksichtigt und selten auch der *systolische Druck*.
Einige der Indikatoren für eine hinreichende Durchblutung basieren auf einer Untersuchung der Extremitäten. Es wird geprüft, ob die *Beine marmoriert*, oder die *Füße kalt* sind. Ein weiterer Test sieht vor, nach einem Druck auf die Fingerkuppe die Zeit zu messen, bis sich die Gefäße wieder mit Blut gefüllt haben.
Zur Bewertung der Atmung wird der *Quotient aus Sauerstoffpartialdruck im arteriellen Blut und dem Anteil Sauerstoff an der Einatemluft* (p_aO_2 / F_iO_2) betrachtet. Weiterhin gehen aus der Blutgasanalyse die *Differenz aus arteriellem und venösem Sauerstoffgehalt* ($C_{a-v}O_2$), der *gemischt-venöse Sauerstoffpartialdruck* (p_vO_2) und die *gemischt-venöse Sauerstoffsättigung* (S_vO_2) in das Protokoll ein.
Im Blut wird weiterhin die Konzentration verschiedener Stoffe gemessen. Diese sind *Creatinin*, *Laktat*, *Natrium* und *Glukose*.
Auch zur Nierenfunktion und zum Flüssigkeitshaushalt werden mehrere Variablen betrachtet. Die *Gesamtflüssigkeitsaufnahme* und der *Urinfluss* werden bilanziert und die *Osmolalität des Urins* wird bestimmt. Zur Kenntnis genommen wird natürlich auch, wenn für einen Patienten, als Folge einer unzureichenden Nierenfunktion, ein *Nierenersatzverfahren* (Dialyse) zum Einsatz kommt. Ein Hinweis auf ein zu hohes Flüssigkeitsvolumen im Kreislauf ist das Vorhandensein von Flüssigkeitsansammlungen im Gewebe (*Ödeme*). Ein für verschiedene Berechnungen benötigter Wert ist das *Körpergewicht*.

Verwendete Interventionen

Das Protokoll sieht als Interventionen lediglich die Gabe von Medikamenten und Infusionen vor. Die folgende Übersicht zeigt die Auswirkungen der verwendeten Wirkstoffe auf die verschiedenen Parameter. In Klammern stehen teilweise die in späteren Experimenten verwendeten Medikamente zu den Wirkstoffen.

[2] Hier spielen Verfügbarkeitsaspekte eine Rolle.

- Adrenalin (Suprarenin)

  In niedriger Dosierung erhöht Adrenalin das Herzzeitvolumen, durch Steigerung der Kontraktionskraft des Herzens und der Herzfrequenz. Dadurch erhöht sich auch der arterielle Blutdruck.

  In mittlerer Dosierung erhöht sich zusätzlich der arterielle Gefäßwiderstand. Der arterielle Blutdruck erhöht sich weiterhin.

  In hoher Dosierung werden überwiegend der arterielle Gefäßwiderstand und der arterielle Blutdruck erhöht.

- Dobutamin (Dobutrex)

  Dobutamin dient der Steigerung der Herzleistung.

  Das Herzzeitvolumen erhöht sich infolge einer Steigerung sowohl der Kontraktionskraft des Herzens, als auch der Herzfrequenz. Dadurch erhöht sich auch der arterielle Blutdruck.

  In hoher Dosierung (ab 10 μg/kg/min) tritt eine Verminderung des arteriellen Gefäßwiderstandes mit eventueller Senkung des arteriellen Blutdrucks ein.

- Dopamin

  In niedriger Dosierung dient Dopamin lediglich der Verbesserung der Nierenfunktion und Urinproduktion.

  In mittlerer Dosierung (zwischen 3 und 10 μg/kg/min) steigert Dopamin die Herzfrequenz und das Herzzeitvolumen.

  In höheren Dosierungen steigert Dopamin darüber hinaus den Gefäßwiderstand und den arteriellen Blutdruck.

- Noradrenalin (Arterenol)

  Noradrenalin ist ein Vasokonstriktor. Die direkte Wirkung ist eine Erhöhung sowohl des arteriellen Gefäßwiderstandes, als auch des arteriellen Blutdrucks.

  Indirekt[3] vermindert Noradrenalin die Herzfrequenz, das Herzzeitvolumen und die Organdurchblutung.

- Furosemid (Lasix)

  Furosemid ist ein Diuretikum, entzieht dem Körper also über die Niere verstärkt Flüssigkeit.

- Phenylephrine

  Phenylephrine ist ein Vasokonstriktor mit einer Wirkungsweise ähnlich der von Adrenalin und Noradrenalin. Der arterielle Blutdruck steigt infolge einer Erhöhung des Gefäßwiderstandes. In der deutschen Intensivmedizin ist dieses Medikament nicht üblich.

Von besonderer Bedeutung für das Protokoll ist die Unterteilung der Medikamente in Vasokonstringenzien, positiv inotrope Medikamente und Diuretika. Diese wurde bis zur

[3]Diese Wirkung tritt infolge einer Gegenregulation des Körpers ein.

Substanz	Dosierung	Wirkung
Adrenalin	0.01 - 0.03 μg/kg/min 0.03 - 0.15 μg/kg/min 0.15 - 0.30 μg/kg/min	positiv inotrop inotrop und Vasokonstriktor Vasokonstriktor
Dobutamin	1 - 30 μg/kg/min	positiv inotrop
Noradrenalin	0.1 - 0.4 μg/kg/min	Vasokonstriktor
Dopamin	0.5 - 3 μg/kg/min 3 - 10 μg/kg/min 10 - 20 μg/kg/min	kein Effekt auf Kreislauf positiv inotrop Vasokonstriktor
Furosemid	jede Dosierung	Diuretikum
Phenylephrine	Unklar. Wird nicht verwendet.	—

Tabelle 3.1.: Zuordnung von Medikamenten zu Vasokonstringenzien, positiv inotropen Medikamenten und Diuretika

vorliegenden Version jedoch erst unvollständig ausgearbeitet. Besonders um die Eigenschaft einiger Medikamente, dosisabhängig unterschiedliche Wirkungen zu zeigen, musste das Protokoll ergänzt werden.
In Zusammenarbeit mit einem medizinischen Experten wurde das Protokoll um einige Zuordnungen bereichert. Tabelle 3.1 gibt das Ergebnis wieder.
Neben Medikamenten schreibt das Protokoll die Gabe verchiedener Infusionen vor. Dies sind neben einer reinen isotonen Kochsalzlösung (Ringer-Lösung) auch noch Infusionen mit 5% Glukose-Zusatz und Infusionen die 5% Glukose mit Kochsalzlösung im Verhältnis 0.25 zu 0.45 mischen, was etwa vergleichbar mit dem hierzulande verwendeten „Tutofusin OPX“ ist.
Die Verwendung von Infusionen, die Mannitol enthalten, ein diuretisch wirkendes Medikament (Osmodiuretikum), wird vom Protokoll zwar einige Stunden lang in den Entscheidungsfindungsprozess mit einbezogen, jedoch sind Verordnungen derartiger Infusionen nicht Bestandteil des Protokolls.

3.3. Vorbemerkungen zu wissensbasierten Systemen

3.3.1. Formalismen und Wissensrepräsentation

Damit ein Rechner Wissen über einen Sachbereich und Problemlösungsstrategien verwenden kann, muss dieses in einer angemessenen Form verfügbar sein. Für die in Abschnitt 2.1 erwähnte Maschinenlesbarkeit des Behandlungsprotokolls ist eine formale Repräsentation zwingend notwendig. Damit das Modell später diskutierbar ist, ist die Übersetzung in eine deklarative Form notwendig.
Wissensrepräsentation behandelt die Darstellung von Wissen anhand verschiedener Formalismen, denen eine formale Semantik, üblicherweise auf Logik basierend, zugeordnet ist. Ein Interpreter operationalisiert den logischen Folgerungsbegriff und vermag so das repräsentierte Wissen zu verarbeiten. Die Ausdrucksstärke der Formalismen auf der einen Seite, und deren Berechenbarkeits- bzw. Komplexitätseigenschaften auf der anderen sind dabei von zentraler Bedeutung. Eine größere Ausdruckskraft des Formalismus geht mit einer größeren Berechenbarkeitskomplexität einher. Eine kurze Einführung in das Forschungsgebiet der Wissensrepräsentation bietet [OKKN00].

Einschränkungen der Prädikatenlogik

Ein Formalismus, der die Repräsentation von Wissen in der Form deklarativer Regeln gestattet, stellt die Prädikatenlogik erster Stufe dar, beispielsweise beschrieben in [NS97]. Der prädikatenlogische Resolutionskalkül ist für diesen Formalismus widerlegungsvollständig und ermöglicht eine Operationalisierung des Folgerungsbegriffs. Auf diesem Prinzip beruht die Programmiersprache PROLOG. Der von PROLOG verwendete Formalismus ist allerdings auf allquantifizierte Hornformeln eingeschränkt. [Bra87] stellt eine umfassende Einführung in PROLOG dar.
Die Ausdrucksstärke der Prädikatenlogik erster Stufe ist zur Umsetzung des Behandlungsprotokolls hinreichend, sofern Vergleiche und arithmetische Operationen über numerischen Domänen unterstützt werden. Allerdings handelt es sich dabei um einen Formalismus, der im allgemeinen Fall keine effiziente Verarbeitung des repräsentierten Wissens erlaubt, da die logische Folgerung nur semientscheidbar ist. Wird die volle Prädikatenlogik auf funktionsfreie Hornlogik eingeschränkt, so lassen sich über Grundfaktentransformationen in endlicher Rechenzeit inferentielle Hüllen ableiten. Auf dieser Tatsache beruht beispielsweise das logische Datenmodell LOGODAT[4]. In einer derart eingeschränkten Prädikatenlogik ist insbesondere die logische Folgerung entscheidbar.

Modellieren mit MOBAL

Die Modellierung geschah in dieser Diplomarbeit mit Unterstützung des Systems MOBAL. In diesem System kann Wissen direkt vom Anwender in einem eigenen Formalismus eingegeben, aber auch über Werkzeuge aufbereitet oder gar automatisch auf der Basis verfügbarer Daten akquiriert werden. Teil des Systems ist ein Interpreter, der Konsequenzen repräsentierten Wissens ableiten und Widersprüche aufzeigen kann. Eine knappe Einführung in MOBAL stellt [ES] dar. Eine umfassendere Beschreibung, die auch theoretische Aspekte berücksichtigt, findet sich in [MWKE93].

[4]Eine Beschreibung von LOGODAT sowie ein Beweis zur Vollständigkeit der Grundfaktentransformation finden sich in [Bis95]

Der von MOBAL verwendete Formalismus basiert auf funktionsfreier Hornlogik, geht in verschiedener Hinsicht jedoch darüber hinaus. Für diese Arbeit kann der Formalismus als vierwertige Logik angesehen werden, bei dem jedem Grundfakt einer der Wahrheitswerte „wahr", „falsch", „widersprüchlich" oder „nichts bekannt" zugeordnet wird. „Widersprüchlich" wird zugeordnet, wenn ein Grundfakt zusammen mit dessen Negation abgeleitet wird. „Nichts bekannt" bedeutet, dass weder das Grundfakt, noch dessen Negation abgeleitet wurde. Intern werden diese Wahrheitswerte über sogenannte Evidenzen repräsentiert. Dies sind Tupel aus der Menge $\{0, 1000\} \times \{0, 1000\}$. Der erste Wert gibt jeweils die Evidenz für das positive, der zweite die Evidenz für das negative Literal an. Ein Wert von 1000 bedeutet, dass das jeweilige Literal (positiv oder negativ) abgeleitet wurde, ein Wert von 0 hingegen, dass nicht. Prinzipiell sind in MOBAL, im Sinne von unscharfer Logik, auch alle dazwischenliegende Werte möglich. Evidenzen können also beliebige Elemente aus der Menge $[0, 1000] \times [0, 1000]$ sein. In dieser Arbeit kommen nur die oben genannten vier Elemente vor.

Die Bedeutung der verschiedenen Evidenzen, und wie Prämissen formuliert werden müssen, um diese gezielt abzufragen, soll ein Beispiel demonstrieren. Nehmen wir an, wir wollen den Begriff „Tachykardie" als „Herzfrequenz größer 140" definieren. An späterer Stelle soll sowohl bekannte Tachykardie, als auch bekannte „Nicht-Tachykardie" aus dem Ergebnis ableitbar sein. Da der Begriff auf der Herzfrequenz basiert, kann dieser nicht abgeleitet werden, wenn die Herzfrequenz unbekannt ist.

Für den Moment sei die Herzfrequenz über das Literal $\mathsf{heartrate(Patient, Time, HR)}$ mit gebundenen Variablen $\mathsf{Patient}$ und Time zu erfragen, woraufhin die Variable HR mit der Herzfrequenz des Patienten zum angegebenen Zeitpunkt gebunden werde. Liegt zum angegebenen Zeitpunkt für den Patienten kein Wert der Herzfrequenz vor, so werde das obige Literal unter keiner Belegung wahr. Das zweistellige Prädikat $\mathsf{tachycardia}$ gebe das Ergebnis der Auswertung wieder, wobei die zwei Stellen den Patienten und den Zeitpunkt, auf den sich die Aussage bezieht angeben. Weiterhin seien die Vergleichsrelationen „$>$" und „$\leq$" verfügbar.

Dann leitet die folgende Regel eine bekannte Tachykardie ab:

$$\begin{array}{ll} & \mathsf{heartrate(Patient, Time, HR)} \\ \wedge & \mathsf{HR} > 140 \\ \longrightarrow & \mathsf{tachycardia(Patient, Time)}. \end{array}$$

Liegt bekannterweise keine Tachykardie vor, so wird dies durch folgende Regel expliziert[5]:

$$\begin{array}{ll} & \mathsf{heartrate(Patient, Time, HR)} \\ \wedge & \mathsf{HR} \leq 140 \\ \longrightarrow & \neg\mathsf{tachycardia(Patient, Time)}. \end{array}$$

Ist die Herzfrequenz unbekannt, so ist keine der Regeln anwendbar, da die erste Prämisse dann unerfüllbar ist. Unbekannte Werte stellen einen eigenen Fall dar. Der Formalismus von MOBAL stellt die Mittel bereit, um diesen gezielt abzufragen. Der eingebaute Operator $\mathsf{unknown}$ oder \+ bekommt ein Literal als Argument übergeben, und wird mit dem Wahrheitswert „wahr" belegt, genau dann, wenn das als Argument übergebene Literal unter keiner Belegung der freien Variablen wahr wird. Damit verhält sich dieser

[5] Die Negation wird hier durch „¬" statt wie in MOBAL-Syntax durch „not(...)" notiert.

Wahrheitswert	Prämisse	Evidenz
wahr	tachycardia(Patient, Time)	[1000, 0]
falsch	¬tachycardia(Patient, Time)	[0, 1000]
„wahr“ nicht ableitbar	\+(tachycardia(Patient, Time))	[0, ?]
„falsch“ nicht ableitbar	\+(¬tachycardia(Patient, Time))	[?, 0]
nichts bekannt	\+(tachycardia(Patient, Time)) ∧ \+(¬tachycardia(Patient, Time))	[0, 0]

Tabelle 3.2.: Übersicht über Evidenzen und zugehöriger Abfragen am Beispiel der Tachykardie

Operator wie die Negation in Prolog, drückt also ein „negation by finite failure“ aus[6]. Bei Verwendung dieses Operators ergibt sich eine nicht-monotone Regelmenge, das heißt Ableitungen können ungültig werden, wenn neue Fakten hinzugefügt werden.
Tabelle 3.2 zeigt, wie Wahrheitswerte (in einem etwas weiteren Sinne) und Evidenzwerte zusammenhängen, und wie sich die einzelnen Abfragen gestalten.
Neben den oben erwähnten Hornformeln, die von wahren oder falschen Grundfakten auf die Wahrheit oder Falschheit anderer Grundfakten schließen, existieren noch sogenannte „Integrity Constraints“. Diese sind syntaktisch den übrigen Regeln ähnlich, dürfen aber mehrere konjunktiv oder disjunktiv verknüpfte Konklusionen besitzen. Sinn der Constraints ist die Verletzung notwendiger Bedingungen des Modells zu verhindern. Falls eine solche Regel unter einer bestimmten Variablenbindung falsch wird, so wird dies durch einen Eintrag auf der sogenannten „Agenda“ vermerkt. Auch widersprüchlich abgeleitete Grundfakten (Evidenz [1000, 1000]) werden auf die Agenda gesetzt. Auf diese Weise können Fehler im Modell zu einem geeignetem Zeitpunkt korrigiert werden. Zur Fehlerkorrektur steht mit KRT ein Werkzeug für minimale Basisrevisionen zur Verfügung. Weiterhin hilfreich ist die Möglichkeit, in MOBAL die Ableitungen von Grundfakten über Graphen zu visualisieren. Auch auf diese Weise lassen sich die Ursachen von Fehlern schnell finden.
Die Unterstützung von Sorten stellt eine weitere hilfreiche Eigenschaft von MOBAL dar. Diese bilden ein geeignetes Mittel, um die verwendete Konstantenmenge zu strukturieren. Sorten können als unverträglich (disjunkt) deklariert werden, oder durch die Angabe einer Obermenge in eine Hierarchie eingefügt werden. Es wird automatisch der bezüglich der Teilmengenrelation entstehende Verband berechnet. Bei der Deklaration von Prädikaten können den einzelnen Stellen Sorten zugeordnet werden. Das Prädikat tachycardia aus dem obigen Beispiel könnte etwa folgenderweise deklariert werden:

tachycardia/2: <patient>, <time>.

Einträge zu dem Prädikat in der Faktenbasis führen zu automatischen Erweiterungen der Sorten um die Konstanten an den entsprechenden Stellen.
So gehört nach der Ableitung oder dem Eintrag von tachycardia(1234, 20) die Konstante 1234 zur Sorte <patient> und 20 zur Sorte <time>. Diese Eigenschaft ist wiederum bei der Suche nach Fehlern hilfreich, da etwa vertauschte Variablen in Regeln oft zu augenscheinlich falschen Zuordnungen von Konstanten zu Sorten führen.

[6]Anders als in PROLOG ist die in Inferenztiefe in MOBAL allerdings beschränkt. Die maximale Inferenztiefe ist ein Parameter der Inferenzmaschine.

Die ARDEN-Syntax

Eine weniger attraktive aber erwähnenswerte Alternative zu MOBAL stellt eine Modellierung mit der ARDEN-Syntax dar. Dabei handelt es sich um einen Standard für medizinische Wissensbasen, der dazu dient, diese übergreifend nutzbar zu machen. [Ma95] gibt einen kurzen Einblick in einige Eigenschaften des Standards. Die in Abschnitt 1.1 bereits kurz erwähnte Wissensbasis, die u.a. der Unterstützung bei Arzneimittelverordnungen dient, basiert beispielsweise auf der ARDEN-Syntax (vgl. [Wai98]).
Besonders die Unterstützung durch diverse Werkzeuge motiviert allerdings die Wahl von MOBAL als Modellierungsumgebung. Dies ist von besonderem Gewicht, da die Operationalisierung des Behandlungsprotokolls erst ein Zwischenziel ist. Es muss möglich sein, Fehler in der Modellierung zu identifizieren und mit geringem Aufwand zu korrigieren. Diese Arbeitsschritte werden in geeigneter Weise von MOBAL unterstützt.

Zur Notation von MOBAL-Regeln

In späteren Abschnitten werden zur Beschreibung eines Modells häufig in MOBAL repräsentierte Regeln abgedruckt. Die dabei verwendete Notation ist weitestgehend, aber nicht vollständig mit der MOBAL-Syntax identisch.
Einige Hinweise zur Notation:

- Alle Regeln sind als implizit allquantifiert zu verstehen.
- Variablenbezeichner sind genau die Termbezeichner, die mit einem Großbuchstaben beginnen.
- Die Negation wird in MOBAL mittels „**not(...)**“ ausgedrückt, die Konjunktion mittels „&“. In den folgenden Abschnitten werden aber, der besseren Lesbarkeit halber, die in der Logik üblichen Symbole „$\neg$“ und „$\wedge$“ verwendet.
- Da ***unknown(...)*** an die Aussage „nichts bekannt über ...“ erinnert (Evidenz [0,0]), wird die alternative Notation *\+(...)* bevorzugt. Diese ist auch für die Negation in PROLOG (s.o.) gebräuchlich und dürfte von daher eher mit der korrekten Semantik assoziiert werden.
- Vergleiche zwischen Zahlen und zwischen nicht numerischen Konstanten[7] sind in MOBAL über eingebaute Prädikate möglich. Dabei steht ne/2 für ungleich und eq/2 für gleich[8]. Für numerische Vergleiche stehen weiterhin lt/2 und le/2 für „$<$“ und „$\leq$“, sowie gt/2 und ge/2 für „$>$“ und „$\geq$“ zur Verfügung. Wiederum motiviert durch die bessere Lesbarkeit werden die mathematischen Vergleichssymbole in Infix-Notation verwendet.
- Integrity Constraints werden durch die Verwendung des Implikationssymbols „$\leadsto$“ syntaktisch von „normalen“ Regeln unterschieden.
- Sorten stehen in spitzen Klammern. Prädikate werden durch nachfolgende Stelligkeit gekennzeichnet, teilweise gefolgt von der Sortendeklaration für die einzelnen Argumente.

[7] Vergleiche zwischen Zahlen und nicht numerischen Konstanten schlagen hingegen immer fehl, also auch etwa „ne(c, 5.0)“!
[8] Genau genommen steht eq/2 für „unifizierbar“.

- In Anlehnung an das Protokoll ist die verwendete Sprache für logische Bezeichner des Modells Englisch.

3.3.2. Grundsätzliches zum Aufbau einer Wissensbasis

Wissensbasierte Systeme

Wissensbasierte Systeme sind bereits seit mehreren Jahrzehnten Forschungsgegenstand der Künstlichen Intelligenz. Mit MYCIN wurde bereits in den 70er Jahren ein bekanntes regelbasiertes medizinisches Diagnosesystem erstellt. Bei der Untersuchung dieses und weiterer Systeme ließen sich Gemeinsamkeiten erkennen, die eine Terminologie zur Beschreibung auf einer abstrakteren Ebene, der sogenannten Wissensebene, motivierten. Newell beschreibt diese Ebene ausführlich in [New82]. Besonders interessant für diesen Abschnitt sind die von Clancey in [Cla85] geprägten Begriffe zur heuristischen Klassifikation und die später von Puppe in [Pup90] und [PGPB96] entwickelten und zusammengetragenen Begriffe zu Problemlösungsmethoden und Wissensarten. Eine historische Übersicht über die Entwicklung der Forschung im Bereich früher Expertensysteme[9], etwa durch die Schaffung von Expertensystemhüllen wie EMYCIN, oder durch eine zunehmend bessere Trennung von Wissensarten findet sich in [Cla86].
Von anderer Software unterscheiden sich wissensbasierte Systeme besonders durch eine Trennung in sachbereichsunabhängige Problemlösungsmethoden und das Wissen über den jeweiligen Sachbereich. Die Wahl einer spezifischen Problemlösungsmethode legt die Repräsentation des verwendeten Wissens und dessen Verarbeitung fest. Das Sachbereichswissen ist also insbesondere vom Kontrollwissen, das beispielsweise die jeweils nächste zu testende Hypothese bestimmt, abgegrenzt. Je stärker die Vorgaben durch die Methode sind, desto leichter fällt die Organisation des Wissens beim Wissenserwerb, also der Erhebung und Formalisierung von Wissen über einen Sachbereich. Voraussetzung dabei ist, dass die Methode sich für die zugedachte Anwendung eignet. Die Eignung von Methoden lässt sich grob durch die Zuordnung einer Aufgabe zu einem *Problemtyp* beschreiben. Dies ist daher die erste Abstraktion, von einer Anwendung her kommend. Puppe erwähnt noch eine davon verschiedene Gliederung von *Problemlösungstypen*. Diese lassen bereits stärker die Anwendbarkeit spezifischer Methoden erkennen, da sie bereits Art und Organisation des Wissens festlegen. In [Cla85] findet sich eine frühe Einteilung einiger Problemtypen in Ober- und Unterprobleme. Auf der obersten Ebene wird dabei zwischen „Klassifikation“ (analytischen Problemen) und „Konstruktion“ (synthetischen Problemen) unterschieden. Im Laufe der Zeit wurden auch andere, oft ähnliche Gliederungen vorgeschlagen. Eine Übersicht findet sich in [Pup90].

Klassifikationsprobleme

Einen im Kontext wissensbasierter Systeme mittlerweile gut verstandenen Problemtyp stellt die Klassifikation dar. Dabei beschreibt das repräsentierte Wissen Zusammenhänge zwischen beobachtbaren, erfragbaren oder ableitbaren Merkmalen und einer Menge möglicher Lösungen. Bevor ein wissensbasiertes System Schlussfolgerungen ziehen kann, müssen für gewöhnlich erst einige Merkmale über eine Dialogkomponente erfragt werden. Direkt erfragbare Merkmale können beispielsweise Symptome eines Patienten sein. Nachdem hinreichend viele Informationen vorliegen, kann eine der dem System bekannten

[9] Puppe definiert Expertensysteme als wissensbasierte Systeme, bei denen das Wissen von Experten stammt.

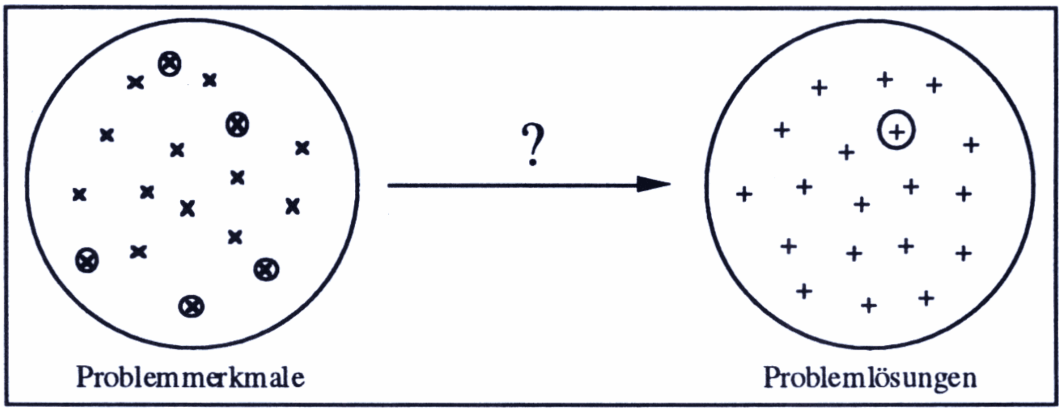

Abbildung 3.1.: Schluss von Merkmalen auf Lösungen, das Wesen von Klassifikationsproblemen. Quelle: [PGPB96]

Lösungen ausgewählt werden. Im medizinischen Kontext kann eine solche Lösung beispielsweise die Form eines Krankheitsbildes haben, das alle erklärungsbedürftigen Symptome zur Konsequenz hat. Schließlich wird die Lösung und deren Herleitung über eine Erklärungskomponente in verständlicher Form ausgegeben.

Definitorisches Merkmal der Aufgabe „Klassifikation“ ist die *Auswahl* einer (oder auch mehrerer) Lösung aus einer vorgefertigten Menge von Lösungen. Im Gegensatz dazu gibt es bei der Problemklasse der „Konstruktion“ keine feste Lösungsmenge, sondern Lösungen müssen aus elementareren Bestandteilen zusammengefügt werden. Das Problem der Planung stellt ein Beispiel dar. Bei der Planung[10] ist das Ziel, eine Folge von Operatoren zu finden, die einen Anfangszustand in einen Zielzustand überführen, wobei für alle Zustände die durch die Operatoren herbeigeführten Veränderungen bekannt sind.

Für den speziell hinsichtlich des Aufbaus einer Wissensbasis besser verstandenen Problemtyp der Klassifikation lassen sich verschiedene Problemlösungstypen angeben, die sich besonders durch die Verwendung unterschiedlicher Wissensarten unterscheiden. Clancey unterscheidet in [Cla85] nur die einfache von der heuristischen Klassifikation. Im nachfolgenden Abschnitt werden anhand verschiedener Wissensarten noch weitere Problemlösungstypen vorgestellt. Bei der *einfachen Klassifikation* findet eine unmittelbare Zuordnung unbekannter Objekte oder Phänomene zu einer bekannten Klasse von Lösungen statt, ohne dass dabei auf unsichere Schlüsse zurückgegriffen wird, oder umfangreiche Zwischenkonzepte in die Herleitung eingehen. Entscheidungsbäume stellen eine mögliche Repräsentation für diese Art der Klassifikation dar. Den inneren Knoten eines solchen Baumes sind Tests zugeordnet, deren Ergebnisse den jeweils zu wählenden Nachfolger bestimmen. Die Blätter geben die zu wählende Lösung an.

Bei der *heuristischen Klassifikation* handelt es sich um einen Problemlösungstyp, der sich durch eine charakteristische Inferenzstruktur beschreiben lässt. Ausgehend von einer Menge von Basismerkmalen werden durch hierarchische Regeln Abstraktionen gewonnen. Solche Regeln können die Form von Definitionen, qualitativen Abstraktionen oder Generalisierungen über Taxonomien besitzen.

[10] Genauer gesagt ist dies das Szenario der klassischen Planung. Neuere Sichtweisen berücksichtigen auch, dass in praktischen Anwendungen die Folgen eines Operatoraufrufs nicht immer vorhersagbar sind. Eine modernere Variante der Planung beschreibt auch [Kli98].

- *Definitionen* sind dabei etwa von der Form „$CI := HZV/KOF$"[11] oder „Von akutem Lungenversagen sprechen wir genau dann, wenn $p_aO_2/F_iO_2 < 200$."
- Bei *qualitativen Abstraktionen* geht es darum, dass numerische Werte oft gar nicht so genau für den Problemlösungsprozess benötigt werden, wie sie erhoben werden, sondern dass die Zusammenfassung vieler ähnlicher Werte diesen sogar erleichtert. Beispielsweise mag es genügen zu wissen, ob der Blutdruck „zu niedrig", „normal" oder „zu hoch" ist, da nur im Falle problematischer Werte interveniert werden soll. Die Abbildung numerischer Werte auf Kategorien, etwa durch Angabe eines festen Intervalls zu jeder Kategorie, stellt eine qualitative Abstraktion dar.
- Taxonomien strukturieren Begriffe bezüglich einer Unterbegriffsrelation. So ist der Begriff „intrathorakaler Blutdruck" ein Oberbegriff zu „rechtsatrialer Blutdruck" und zu „pulmonalkapillärer Verschlussdruck" (vgl. Abschnitt 1.4). Ein einfaches Beispiel für *Generalisierungen über Taxonomien* ist, aus dem Ereignis „pulmonalkapillärer Verschlussdruck verfügbar" das Ereignis „ein intrathorakaler Blutdruck verfügbar" abzuleiten.

Neben der Hierarchie zur Merkmalsvorverarbeitung gibt es noch eine Lösungshierarchie. Abstrakte Lösungsklassen werden zu den endgültigen Lösungen verfeinert. Der Abstraktionsgrad dieser Hierarchie nimmt also entlang der Inferenzrichtung ab.
Ein besonderes Merkmal der heuristischen Klassifikation ist, dass unsicheres Wissen, in Form von heuristischen Assoziationen, die beiden beschriebenen Hierarchien verbindet. Unterschiedlich abstrakte Merkmale werden also durch unsichere Schlüsse unterschiedlich abstrakten Lösungen zugeordnet.
In Anlehnung an [PGPB96], seien im folgenden *Symptome* die unmittelbar vorliegenden oder erfragten Merkmale, und *Symptominterpretationen* Merkmale, die durch Abstraktion aus Symptomen hergeleitet werden. Lösungsklassen werden als *Grobdiagnosen*, die Menge der endgültigen Lösungen als *Feindiagnosen* bezeichnet. Zum Begriff der *Diagnose* ist anzumerken, dass er in dieser Arbeit synonym zum Begriff „Klassifikation" verwendet[12], aber zur Vermeidung von Verwechslungen mit dem medizinischen Diagnosebegriff weitestgehend vermieden wird.

Wissensarten

Wissensbasierte Systeme verwenden zur Problemlösung unterschiedliche Arten von Wissen. Die bereits erwähnten Abstraktionen, also die Herleitung von Symptominterpretationen aus Symptomen, basieren auf sicherem Wissen und können von den verschiedensten Problemlösern verwendet werden. Die im folgenden erwähnten Wissensarten erfordern jedoch unterschiedliche Methoden. Da für die Methoden die Wissensarten zentral sind, werden sie nach den Wissensarten benannt. In diesem Abschnitt sollen kurz die wichtigsten in [PGPB96] vorgestellten Wissensarten für Diagnosesysteme skizziert, und besonders der bereits beschriebene Problemlösungstyp der heuristischen Klassifikation von den übrigen abgegrenzt werden.
In der heuristischen Klassifikation wird anhand von heuristischen Assoziationen von einer Hierarchie zur Merkmalsvorverarbeitung auf eine Hierarchie von Grob- und Feindiagnosen geschlossen.

[11] Dies ist die Definition des Herzindex (CI). Dabei bezeichnet HZV das Herzzeitvolumen und KOF die Körperoberfläche.
[12] Der Begriff „Diagnose" wird in der Literatur nicht einheitlich gebraucht.

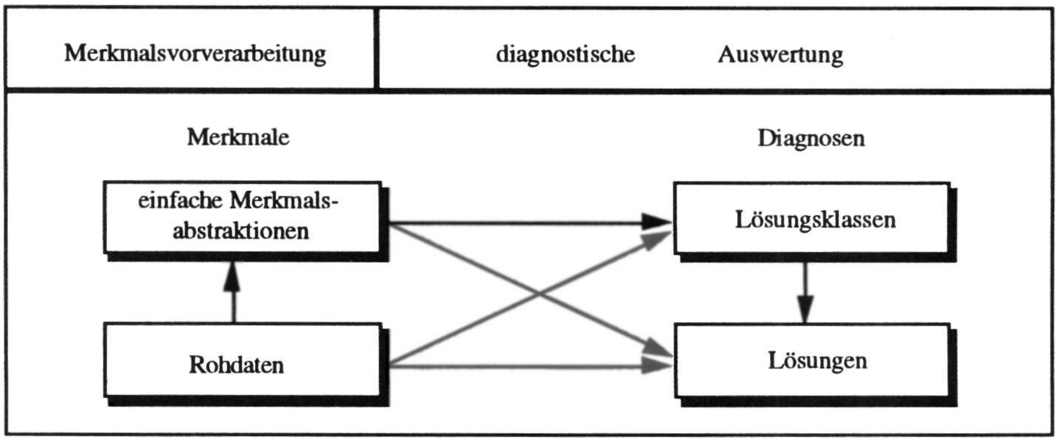

Abbildung 3.2.: Die charakteristische Inferenzstruktur der heuristischen Klassifikation. Quelle: [PGPB96]

Heuristisches Wissen hat oft den Charakter von Faustregeln. Laut [Cla85] beschreiben heuristische Assoziationen manchmal schlecht verstandene Korrelationen, und sie beruhen oft auf Annahmen, dass bestimmte Sachverhalte typisch seien. Aus der Sichtweise der Modellierung eines Sachbereichs überspringen heuristische Assoziationen oft Zwischenkonzepte, beispielsweise indem kausale Zusammenhänge zwischen verbundenen Entitäten nicht in den Regeln expliziert werden. Aus Sicht der Problemlösung stellen sich diese „Abkürzungen" als nützliches Mittel zur Effizienzsteigerung heraus. Heuristische Assoziationen werden daher auch als *kompiliertes Wissen* bezeichnet. Als formales Merkmal können für die Regeln Evidenzwerte angegeben werden. Experten fällt die unmittelbare Angabe solch annotierter Regeln jedoch gewöhnlich schwerer, als einzelne Fälle oder kausale Zusammenhänge anzugeben. Eine Methode, die aus leichter zu erhebendem Wissen effizientere heuristische Regeln ableitet wird in [BGGP93] vorgestellt. Der Nutzen solcher Regeln ist nicht auf die heuristische Klassifikation beschränkt, sondern kann auch im Rahmen anderer Problemlösungsmethoden zur effizienten Hypothesengenerierung eingesetzt werden.

Der Unterschied zwischen heuristischer und *statistischer Klassifikation* besteht darin, dass heuristisches Wissen von Experten geschätzt, statistisches Wissen hingegen über die Auswertung einer Falldatenbank automatisch berechnet wird. Für letzteres ist das Theorem von Bayes zentral und die Falldatenbank sollte den Annahmen des Theorems gerecht werden.

Fallbasierte Klassifikation setzt hingegen auf ein Ähnlichkeitsmaß zur Bestimmung von Lösungen. Um anhand von Daten aus einer Falldatenbank einen neuen Fall klassifizieren zu können, wird neben den Fällen selbst (*Fallwissen*) auch noch *fallvergleichendes Zusatzwissen* für die Definition des Ähnlichkeitsmaßes benötigt. Zur Klassifikation eines neuen Falles werden dann die ähnlichsten Fälle der Falldatenbank herausgesucht und ausgewertet.

Schließlich basieren einige Problemlösungstypen auf einem Modell, das umfassend die für die Aufgabe relevanten kausalen Zusammenhänge und Gesetzmäßigkeiten beschreibt. Ein Nachteil dabei, beispielsweise gegenüber der heuristischen Klassifikation, ist der höhere Rechenaufwand. *Überdeckende Klassifikation* schließt von Diagnosen auf Symptome. Die

Richtung der Inferenzen ist also der des heuristischen Wissens entgegengesetzt. Es wird nach einer oder mehreren Diagnosen gesucht, die die beobachteten Symptome erklären. Die Suche kann sich auch an einem *Fehlermodell* orientieren, das das Verhalten des Systems im Falle bestimmter Fehler beschreibt. Besonders bei technischen Anwendungen, etwa der Identifikation defekter Bauteile einer Schaltung, kann die separate Angabe eines *Strukturmodells* hilfreich sein (vgl. [Str00]). In der medizinischen Domäne ist eine Klassifikation anhand von Fehler- und Strukturmodellen problematischer, da viele Zusammenhänge sehr komplex und noch nicht hinreichend gut verstanden sind.

Problemlösungsmethoden

Wie das in einer Wissensbasis formalisierte Wissen zur Lösung eines Problems verwendet wird, hängt von der Wahl einer spezifischen Problemlösungsmethode ab. Diese legt fest, ob eine bestimmte Präferenzordnung über der Lösungs- oder Regelmenge verwendet wird, ob bestimmte Lösungen fokussiert, also gezielt überprüft werden, wann welche Daten über die Dialogkomponente erfragt werden, usw.

Schwache Problemlösungsmethoden sind universell einsetzbar und können daher unabhängig vom verwendeten Wissen beschrieben werden. Sie setzen wenig bezüglich der Form des verarbeiteten Wissens voraus. Obwohl dies auf den ersten Blick wie ein Vorteil erscheinen mag, sind stärkere Vorgaben bei einem bekannten Problemlösungstyp für den Wissenserwerb günstiger. Hinzu kommt, dass bei schwachen Problemlösungsmethoden das Kontrollwissen für die Reihenfolge der Regelanwendungen etc. nicht deutlich vom Bereichswissen getrennt werden kann, sondern eine formal homogene Regelmenge entsteht. Können hingegen viele Annahmen über einen Problemlösungstyp gemacht werden, so können speziell für diesen Zweck entsprechend effizientere Problemlösungsmethoden entwickelt werden.

Zu den schwachen Problemlösungsmethoden im Bereich der Klassifikation zählen neben Vorwärtsverkettung und Rückwärtsverkettung noch „Hypothesize-and-Test“ und „Establish-Refine“. *„Hypothesize-and-Test“* beschreibt ein Vorgehen, bei dem anhand der vorliegenden Merkmale Hypothesen generiert werden. Zur Überprüfung der Hypothesen werden weitere Merkmale angefordert, die dann zur Generierung besserer Hypothesen verwendet werden, usw. *„Establish-Refine“* stellt einen Spezialfall dar. Diagnosen sind dabei über eine Verfeinerungsrelation geordnet, und ausgehend von Grobdiagnosen werden jeweils Spezialisierungen dieser Lösung fokussiert.

Für jeden Problemlösungstyp existieren eigene Problemlösungsmethoden. Da dies im Rahmen dieser Arbeit von untergeordneter Bedeutung ist, soll lediglich noch kurz auf die heuristische Klassifikation eingegangen werden. In [Pup90] und [PGPB96] werden Methoden vorgestellt, die heuristische Assoziationen mit annotierten Unsicherheitswerten verarbeiten. Die Angabe solcher Werte ermöglicht, dass sich ausreichende Evidenz eventuell erst durch das Aufsammeln von Evidenzen, nach mehrfacher Anwendbarkeit entsprechender Regeln ergibt. Ist die Evidenz für/wider eine Hypothese groß genug, so gilt sie als abgeleitet/widerlegt. Bestätigte Hypothesen werden wie Merkmale verwendet. Außerdem können auch mäßig hohe Evidenzen zum Finden einer Lösung beitragen, indem entsprechende Hypothesen zunächst „verdächtigt“ und dann gezielt untersucht (fokussiert) werden.

3.4. Wahl einer geeigneten Modellarchitektur

Um eine geeignete Umsetzung des Protokolls vornehmen zu können, empfiehlt sich zunächst eine Analyse des verwendeten Wissens und der verwendeten Problemlösungsstrategie. Ziel ist nicht die Neuentwicklung eines wissensbasierten Systems, sondern die Umsetzung eines bestehenden, nicht operationalen Protokolls zur Validierung. Aus diesem Grund übertragen sich zwangsläufig einige Probleme aus dem Protokoll auf das zu erstellende Modell. Die Analyse hat die Erstellung einer abstrakteren Beschreibung des Protokolls zum Ziel. Die im Abschnitt 3.3.2 vorgestellte Terminologie ist hierbei hilfreich, da sie ermöglicht, das Systemverhalten unabhängig von Implementationsdetails und der Anwendungsdomäne zu betrachten, was für die Übertragung des Behandlungsprotokolls in eine andere Repräsentation wichtige Anhaltspunkte liefert. Die zentrale Frage in diesem Abschnitt lautet: „Was tut und was weiß das Protokoll?“
Anschließend ist eine Zuordnung zu einem möglichst spezifischen Problemlösungstyp ratsam, um die Formalisierung des Wissens durch die vorgegebene Struktur zu vereinfachen. Ein praktisches Problem ist zunächst einmal nicht unbedingt eindeutig einem Problemlösungstyp zuzuordnen. Lässt es sich unterschiedlichen Problem- und Problemlösungstypen zuordnen, so dürften sich allerdings im Laufe der Modellierung einige als mehr, andere als weniger angemessen erweisen. Entsprechend ist für die im Protokoll behandelte medizinische Aufgabe sowohl der Problemtyp Klassifikation (*Auswahl* einer Therapie), als auch Konstruktion (*Konstruktion* eines Therapieplans unter Rahmenbedingungen) denkbar.[13]
Das vorliegende Protokoll gibt nun jedoch – über das medizinische Problem hinaus – Bedingungen an die Modellierung vor.
Eine Zuordnung des Protokollverhaltens zum Problemtyp der Planung scheitert. Es werden ausschließlich Interventionen für die Gegenwart abgeleitet, während bei der Planung eine sequentielle Operatoren*folge* gesucht wird, die vom Ausgangzustand zu einem Zielzustand führt. Letzteres entspräche etwa einem langfristigen Therapieplan mit mehreren aufeinanderfolgenden Therapien, dem ein vorausschauendes Abwägen der Konsequenzen vieler Alternativen vorausgeht. Die Wahl von Medikamenten erfolgt im Protokoll unmittelbar auf der Basis abstrahierter Patientenvariablen, wobei Zusammenhänge über die Wirkung der Interventionen oft nicht expliziert werden. Zum Verständnis des Protokollverhaltens ist also zusätzliches Expertenwissen notwendig. Dies spricht deutlich gegen die Zuordnung zum Problemtyp der Planung. Weiterhin weist es darauf hin, dass die Vorhersage von Ereignissen oder gar Simulationen[14] anhand des explizierten Wissens nicht möglich ist.
Gegen die überdeckende Klassifikation als Problemlösungstyp spricht, neben der Inferenzrichtung des vorliegenden Protokolls, dass wohl ein wichtiger Teil des für diese Modellierungsform notwendigen Wissens nicht dem Protokoll zu entnehmen ist. So sind, wie bereits erwähnt, die Gründe für die Wahl von Therapien oft nicht dargelegt. Auch Wissen, das für ein Fehler- oder Strukturmodell benötigt würde, geht weit über das Protokoll hinaus. Statistische und fallbasierte Klassifikation widersprechen klar den Zielen dieser Arbeit, da sie aufgrund ihrer Orientierung an Fällen sich weder zur Umsetzung des Protokolls eignen, noch überhaupt ein explizites Behandlungsprotokoll im Sinne von Abschnitt 1.2 beschreiben.
Bei einer Analyse der durch das Protokoll beschriebenen Ableitungen von Wissen lassen sich hingegen, für jede einzelne Visite, die typischen Grundzüge heuristischer Klassifika-

[13] Der Gedanke, die Aufgabe als Planungsproblem aufzufassen, wird in Kapitel 5 nochmal aufgegriffen.
[14] In [Pup90] ist Simulation ein allgemeinster Problemlösungstyp, wie Klassifikation und Konstruktion

tion erkennen. An die für diesen Problemlösungstyp charakteristische Inferenzstruktur erinnert zunächst besonders die Merkmalsvorverarbeitung des Protokolls. Neben direkt vom Arzt erhobenen Variablen gibt es Datenabstraktionsschritte. Die Herleitung von Interventionen geschieht im Protokoll über abstraktere Zwischenschritte. So wird zunächst oft die Erhöhung eines Medikaments aus einer Medikamentengruppe, etwa der Vasokonstringenzien, abgeleitet. Erst im nächsten Schritt wird dann das konkrete Medikament gewählt und schließlich die neue Dosis berechnet. Dies spricht im übrigen auch gegen die einfache Klassifikation als Problemlösungstyp, da diese aufgrund der fehlenden Zwischenkonzepte bei der Verarbeitung (besonders bezüglich der Datenabstraktion) zu ausdrucksschwach ist. Die Verfeinerung zunächst grob abgeleiteter Lösungen grenzt allerdings bereits an die *Konstruktion* von Lösungen, da dabei die neuen Dosierungen von Medikamenten berechnet werden. Die Interventionsvorschläge in leicht vereinfachter Form, beispielsweise „erhöhe die Dosis von Dopamin um 3 μg/kg/min“, stellen jedoch noch eine feste Menge vorbereiteter Lösungen dar, da die jeweiligen Dosisänderungen sich auf nur wenige Möglichkeiten beschränken. Sind die neuen Dosierungen Teil der Lösung, so gibt es derart viele Lösungen, dass diese intuitiv weniger ausgewählt, sondern eher konstruiert werden. Außerdem nicht ganz konform mit den Grundzügen der heuristischen Klassifikation ist die Möglichkeit, dass mehrere Interventionen innerhalb einer Visite ableitbar sind. Beide Argumente sollen in diesem Rahmen allerdings nicht überbewertet werden, und seien eher der Vollständigkeit halber erwähnt[15].
Diese erste Analyse des Protokollverhaltens legt für eine Modellierung die folgende, noch recht grobe Abbildung auf das durch die heuristische Klassifikation vorgegebene „Inferenzgerüst“ nahe:

- Die Eingaben des Protokolls bilden die Symptome.
- Merkmale, die durch Abstraktionsschritte aus Symptomen abgeleitet werden, bilden die Symptominterpretationen.
- Zu abstrakteren „Therapien“ zusammengefasste Interventionen entsprechen den Grobdiagnosen.
- Konkrete Interventionsvorschläge bilden die Feindiagnosen.

Die Problembeschreibungen werden also bezüglich durchzuführender Interventionen klassifiziert. Nach dieser ersten groben Abbildung auf einen Problemlösungstyp ist nun eine feinere Strukturierung des Protokolls notwendig.
Zu den im Protokoll nicht unmittelbar identifizierbaren Bestandteilen gehören Grobdiagnosen. Die vom Protokoll verwendeten Interventionen lassen sich zwar zu wenigen allgemeineren „Therapien“ zusammenfassen, jedoch verhinderten einige Protokollfragmente zunächst eine unmittelbare Umsetzung hierarchisch angelegter Verfeinerungen. Zunächst waren daher einige Umstrukturierungsmaßnahmen[16] notwendig. Hierzu gehörte beispielsweise, die Konfliktauflösung bei mehreren widersprüchlichen Therapiekandidaten zusammenfassend zu behandeln, und im inferentiellen Ablauf bereits auf Grobdiagnosen zu

[15] Die Grenze zwischen heuristischer Klassifikation und heuristischer Konstruktion, wie sie in [Cla85] verstanden werden, ist fließend. Unter heuristische Konstruktion fällt beispielsweise die Auswahl vielversprechender Lösungsfragmente wie bei der heuristischen Klassifikation, wobei erst in einem darauf folgenden Schritt aus den Fragmenten eine Lösung konstruiert wird.

[16] Das Verhalten des Protokolls wurde durch solche Maßnahmen nicht verändert.

beziehen. Das Protokoll leitet hingegen bei unvereinbaren Grobdiagnosen erst noch Feindiagnosen her, die erst unmittelbar vor der Durchführung verworfen werden. Zum Begriff der Diagnose sei nochmal ausdrücklich betont, dass das Klassifikationsziel keine Diagnose im medizinischen Sinne ist, sondern ein Interventionsvorschlag. In abstrakterer Form („erhöhe die Dosierung eines positiv inotropen Medikaments“) werden diese als Grobdiagnosen bezeichnet. Die konkreten Interventionsvorschläge („Dobutamin auf 5 μg/kg/min heraufsetzen“) bilden die Feindiagnosen.
Leichter zu strukturieren sind die Merkmalsabstraktionen des Protokolls. Aus den Symptomen werden durch Definitionen und qualitative Abstraktionen die notwendigen Begriffe hergeleitet. In [PSS00] wird auf die wichtige Rolle der Fachsprachen bei der Modellierung eines Sachbereichs eingegangen. In [PGPB96] (S. 164) propagiert Puppe die Verwendung der Fachterminologie, speziell auch bei medizinischen Wissensbasen. Durch die Verwendung eines „diagnostischen Mittelbaus“, also von Symptominterpretation und Grobdiagnosen, werden sowohl eine höhere Übersichtlichkeit, als auch eine geringere Komplexität und höhere Modularität erzielt. Im Sinne dieser Gedanken wurde das fachbegriffliche Gerüst des Protokolls durch eigene Entitäten in das Modell übertragen. Der Abstraktionsschritt, also die Ableitung von Symptominterpretationen, bedeutet hauptsächlich, dass der Patientenzustand in fachsprachlichen Begriffen ausgedrückt wird. Auf diese Begriffe aufbauend können dann die Regeln zur Behandlung formuliert werden. Natürlich ist dabei zu beachten, dass die Begriffe der Fachsprache in diesem Kontext exakt in ihrer formal definierten, nicht etwa einer abweichend intendierten Semantik, zu verstehen sind. Die formale Standardisierung von medizinischen Begriffen ist aber laut Morris (siehe Abschnitt 1.2) gerade als ein Vorteil expliziter Modelle anzusehen.
Die heuristischen Assoziationen des Protokolls verbinden die Symptome und Symptominterpretationen, die den Patientenzustand charakterisieren, mit Therapiekandidaten. An dieser Stelle bestehen die meisten Freiheitsgrade für ein Behandlungsprotokoll. Als Beispiel soll ein zu niedriger arterieller Blutdruck dienen. Um diesem Problem zu begegnen können drei verschiedene Wege eingeschlagen werden:

1. Es kann eine schnell laufende Infusion verordnet werden. Das zusätzliche Volumen im Kreislaufsystem führt zu einem höheren Druck.

2. Die Schlagkraft des Herzens kann durch ein entsprechendes Medikament erhöht werden. Das Blut wird dann mit höherem Druck in den Körperkreislauf gepumpt.

3. Durch die Gabe von Vasokonstringenzien verengen sich Blutgefäße, was aufgrund des erhöhten Gefäßwiderstands zu einem höheren Druck führt.

Im Sinne des Abschnitts 1.2, über die Standardisierung in der Medizin, ist eine der Leistungen des Protokolls eben die Ausräumung der Freiheitsgrade bei der Therapiewahl. Die Entscheidung für eine dieser Möglichkeiten geschieht daher im Protokoll mit – rein formal betrachtet – sicheren Schlüssen. Dennoch lassen sich diese Regeln von ihrer Fundierung her auch gut als Heuristiken interpretieren, auch wenn dabei keine Unsicherheitswerte Verwendung finden. Wie sich später noch herausstellen wird, fällt die Wahl einer Therapie regional unterschiedlich aus. Dass die Behandlungen sich unterscheiden ist ein Indiz für nicht explizit gemachte Unsicherheit einiger Schlüsse. Dies gilt allerdings nur, solange alle Regeln als Wissen aufgefasst werden. Eine alternative Sichtweise wäre, dass jedes Behandlungsprotokoll einfach eines von vielen möglichen Verhalten beschreibt, das aber an den zweifelhaften Stellen eben nicht unbedingt auf Wissen gegründet

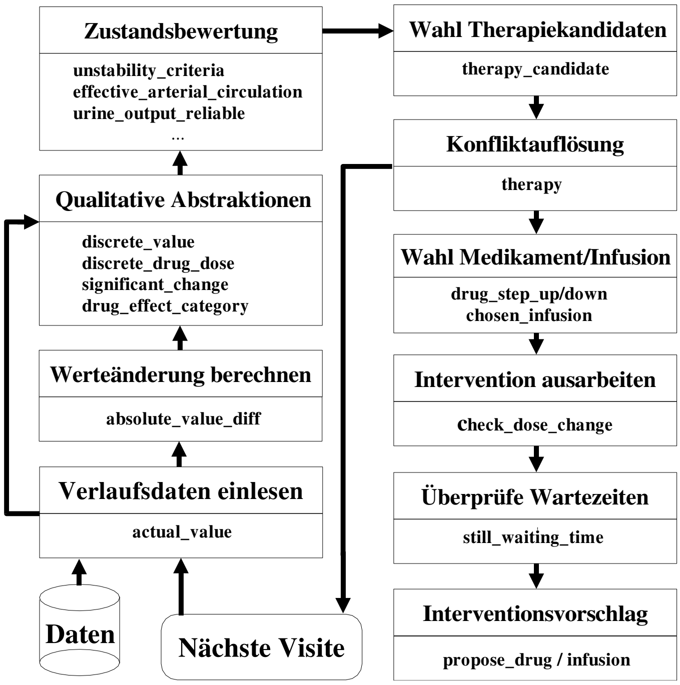

Abbildung 3.3.: Architektur der Wissensbasis

ist. Die Formalisierung dient in einem solchen Kontext der Identifikation eines erfolgreichen Protokolls, beispielsweise durch klinische Studien (vgl. Abschnitt 1.2). Für die hier durchgeführte Modellierung lässt sich, unabhängig von der eingenommen Perspektive, die Therapiewahl als unsicheres Wissen deuten. Aber selbst ohne den heuristischen Aspekt der Therapiewahl ist es sinnvoll, an dieser Stelle eine Verbindung zwischen zwei Hierarchien herzustellen. Die Abstraktionen der Patientenvariablen und die Spezialisierung der Interventionen bilden auf natürliche Weise eigene Hierarchien. Auch Clancey propagiert in [Cla85] am Beispiel des Systems GRUNDY eine saubere Trennung unterschiedlicher Hierarchien. Eine solche Strukturierung des Modells ist übersichtlicher und kommt der Änderbarkeit sehr entgegen, etwa um Experimente mit dem Protokoll zu ermöglichen.

3.5. Verfeinerung der Schablone der heuristischen Klassifikation

In Abschnitt 3.4 wurde anhand einer Analyse des in dieser Diplomarbeit verwendeten Behandlungsprotokolls eine Architektur zu dessen Umsetzung in eine operationale Form gewählt. Die heuristische Klassifikation definiert einen Rahmen, in den das sachbereichsabhängige Wissen nun eingefügt werden muss. Dieser Rahmen wird maßgeblich durch die inferentielle Struktur des Wissens beschrieben.
Während die heuristische Klassifikation noch eine recht grobe, sachbereichsunabhängige Struktur vorgibt, beschreibt die Skizze 3.3 bereits das Ergebnis nach weiteren Schritten. Die feinere Architektur der inferentiellen Struktur spiegelt spezifische Eigenschaften des Protokolls wider, und eignet sich, aufgrund des Detaillierungsgrades, bereits als Schablone, in die das zu repräsentierende Wissen eingetragen werden kann. Die Aggregation von

Begriffen, also das Bilden von Kategorien[17] liefert eine feinere Strukturierung und hilft dadurch, die Übersichtlichkeit und Modularität zu erhöhen.
Die Ausarbeitung dieser Architektur geschah hauptsächlich anhand von über die in Abschnitt 3.4 hinausgehende Analysen des Behandlungsprotokolls. Dies führte jedoch nicht gleich im ersten Anlauf zu einem befriedigenden Ergebnis. Die Bildung ungeeigneter begrifflicher Aggregationen stellt die Hauptschwäche früherer Modellierungsansätze dar. So schien es zunächst, als gäbe es im Protokoll eine natürliche Unterscheidung zwischen der Abstraktion von Medikationen und der, der übrigen Patientenvariablen. Da die Medikationen auf den ersten Blick keine Symptome im medizinischen Sinne darstellen, schienen zwei getrennte Abstraktions-Hierarchien eine intuitiv sinnvolle Art der Modellierung zu beschreiben. Es ergab sich jedoch bei fortschreitender Modellierung Bedarf für eine Zwischenkategorie, da einige Begriffe intuitiv den Symptominterpretationen zuzurechnen sind, aber sowohl auf Variablen zur Medikation, als auch auf anderen Variablen beruhen. Ein bei der Therapiewahl häufig verwendeter Begriff zur Verlässlichkeit des Urinflusses beruht beispielsweise nicht nur auf den im Blut gemessenen Konzentrationen von Glukose und Laktat, sondern auch darauf, ob in den letzten Stunden diuretische Medikamente verabreicht wurden. Nach genauerem Hinsehen erwies sich die Modellierung als unangemessen. Die Medikation eines Patienten in der (näheren) Vergangenheit und Gegenwart geht im Protokoll als ein Teil des aktuellen Patientenzustands in die Therapiewahl ein. Dies wird besonders dort deutlich, wo die Medikation als Indikator für die Stabilität eines Patienten fungiert. Ein stabiler Blutdruck nach mehreren Stunden ohne medikamentöse Unterstützung ist anders zu bewerten, als der gleiche Blutdruck nach massiver Gabe von Vasokonstringenzien. Da die Abstraktion der Medikationen sogar ähnliche Stadien durchläuft, wie die Abstraktion der übrigen Verlaufsdaten, erwies sich die einfachere Architektur aus Skizze 3.3 letztlich als geeigneter.
Dieses Beispiel soll illustrieren, welche Schwierigkeit auf dem Weg zu der recht einfach anmutenden Architektur zu meistern waren. Negative Konsequenzen frühzeitig gefällter Entscheidungen zur Modellierung stellten sich oft erst bei vorangeschrittener Modellierung heraus. Dies reichte von der Definition der Schnittstelle zu den Verlaufsdaten bis hin zu der Frage, ob die Interventionsformen „Infusionen" und „Medikamente" über die gleichen Kategorien abgeleitet werden sollen. Oftmals war es günstiger[18], eine relativ einfache Struktur des Modells zu wählen, und notfalls durch kleine verhaltenskonforme Anpassungen des Protokolls viele Fälle mit wenigen Regeln abzudecken.
Da das umgesetzte Protokoll im Rahmen dieser Arbeit nicht veröffentlicht wird, bietet sich für die weiteren Ausführungen eher die Präsentation des Modellierungsergebnisses an, statt den schrittweisen Aufbau zu beschreiben. In Abschnitt 3.8.2 werden dann einige bei der Umsetzung aufgetretene Probleme diskutiert.

[17] Die Kategorien entsprechen den Kästen in der Skizze. Die Beschriftungen sind allerdings eher Illustrationen zu den Ableitungen der in den Kategorien enthaltenen Begriffe.
[18] Das Maß für „günstig" ergibt sich aus den Anforderungen Übersichtlichkeit, möglichst geringe Komplexität und Modularität.

3.6. Beschreibung der Wissensbasis

In diesem Abschnitt wird das entwickelte Modell vorgestellt, das eine operationale Form des verwendeten Behandlungsprotokolls darstellt. Die Gliederung ist dabei eng an die in Abbildung 3.3 verwendeten Begriffe angelehnt.

3.6.1. Datenschnittstelle

Der Zugriff auf die benötigten Verlaufsdaten geschieht über eine Schnittstelle, die mit Hilfe externer Prädikate realisiert ist. Positive Literale dieser externen Prädikate dürfen in Regel-Prämissen auftreten. Technische Details zu der Schnittstelle finden sich in Anhang A.

Die Verwendung einer „externen Prämisse“ ist vergleichbar mit der Anfrage über eine Dialogkomponente in Diagnose-Systemen, allerdings werden die Anfragen über einer Datenbank ausgewertet. Die Auswertung von Anfragen ist gegenüber dem in MOBAL beschriebenen Modell gekapselt. Welche Anfragen über die Schnittstelle ermöglicht werden entscheidet letztlich auch darüber, welche Vorverarbeitungsschritte der Daten im MOBAL-Formalismus repräsentiert werden, und welche von einem „externen Dienst“ erbracht werden. Für eine Repräsentation im MOBAL-Formalismus sprechen die Möglichkeit zur graphischen Aufbereitung von Ableitungen, das Werkzeug KRT für minimale Basisrevisionen und weitere in MOBAL verfügbare Hilfsmittel. Bei einigen Aspekten des Modells ist die Verarbeitung in MOBAL allerdings erfahrungsgemäß recht ineffizient. Sind dies gleichzeitig Aspekte, die sich auf den Anfang der Inferenzkette beschränken, die sich gut kapseln lassen, denen eine wohldefinierte Semantik zugeordnet werden kann und die aufgrund ihrer elementaren Struktur keinerlei Revisionen benötigen, so erweist sich eine Auslagerung als günstig.

Die im folgenden beschriebenen Prädikate bilden den Großteil der Schnittstelle zur Datenbank. Es fehlen lediglich einige eher technisch motivierte Prädikate. Die Notation ist an PROLOG-Konventionen angelehnt. Variablen, die nicht gebunden sein müssen, werden durch vorangestelltes „-“ gekennzeichnet und werden bei der Auswertung an einen Grundterm gebunden. Die durch ein vorangestelltes „+“ gekennzeichneten Variablen müssen hingegen bereits vor der Auswertung der Anfrage gebunden sein[19]. Die Auswertung einer Anfrage kann auch fehlschlagen. Dies äußert sich in gleicher Weise, wie wenn für ein „gewöhnliches Literal“ einer Prämisse keine erfüllende Belegung der ungebundenen Variablen existiert. Die betrachtete Regel ist in diesem Fall nicht anwendbar.

Die Semantik der Datenbank-Anfragen wird in diesem Abschnitt hauptsächlich mit Hilfe von SQL-Ausdrücken beschrieben. Bezüglich des Datenbankschemas wird dabei angenommen, dass die Daten in einer einzelnen Tabelle namens „data“ abgelegt sind. Ein Tupel fasst alle aktuellen Werte zu den Variablen eines erfassten Patienten/Zeitpunkt-Paares zusammen. Das Attribut „id“ diene der Identifikation des jeweiligen Patienten über eindeutige Bezeichner. Das Attribut „minute“ gebe die verstrichene Zeit seit dem Anfang der Datenerfassung an. Das Attribut „Attribute“ steht in den folgenden SQL-Ausdrücken für das in der jeweiligen Anfrage zu spezifizierende „+Attribute“, kann also eine Patientenvariable, ein Medikament oder eine Infusion bezeichnen.

[19] Hier stellt sich die Frage nach der Ordnung der Literale einer Klausel. In der Logik werden Klauseln üblicherweise als Mengen von Literalen aufgefasst. Die Literale in Regel-Prämissen sind gemäß dieser Sichtweise ungeordnet. In MOBAL geschieht die Variablenbelegung in der Reihenfolge, in der die Literale in der Prämisse aufgeführt sind, analog zu PROLOG.

- ask_value(+Patient, +Timestamp, +Attribute, -Value)

  Dieses Prädikat dient dazu, den für einen Patienten und Zeitpunkt eingetragenen Wert eines Attributs aus der Datenbank zu ermitteln. Ist der Wert eines Attributs für den angegebenen Zeitpunkt nicht verfügbar, so schlägt der Aufruf fehl.

  Intern wird die Anfrage sinngemäß in folgenden SQL-Ausdruck übersetzt:

  ```
  SELECT attribute FROM data WHERE id=Patient and minute=Timestamp;
  ```

- ask_latest_value(+Patient, +Time, +Attribute, +ValidInterval, -Value, -ValueFrom)

  Dies beschreibt den häufigsten Anfragetyp an die Datenbank. Für einen Patienten soll – relativ zum Visitenzeitpunkt – der aktuellste Wert einer Variablen ermittelt werden. Als zusätzlichen Parameter akzeptiert dieses Anfrage-Prädikat ein maximales Alter des zurückgelieferten Wertes. Dies hat neben der Vereinfachung der Werteermittlung auch den Hintergrund, dass fast alle auf diese Weise ermittelten Werte höchstens einige Stunden alt sein dürfen. Anfragen an die Datenbank werden effizienter, wenn sie auf kurze Zeitintervalle beschränkt werden können. Ist die Angabe eines zeitlichen Intervalls nicht erwünscht, so kann statt dessen die Konstante unlimited verwendet werden. Liegt im angegebenen Zeitintervall zu der Variablen kein Wert vor, so schlägt der Aufruf fehl. Andernfalls werden der aktuellste Wert und dessen Erhebungszeitpunkt zurückgegeben.

  Die intern generierte Anfrage entspricht folgendem SQL-Ausdruck[20]:

  ```
  SELECT minute, Attribute FROM data
      WHERE (id = Patient)
      AND (minute ≤ Time)
      AND (minute > Time - ValidInterval)
      AND NOT (Attribute IS NULL)
  ORDER BY minute;
  ```

  Ist das Ergebnis der SQL-Anfrage nicht leer, so ergeben sich die Rückgabewerte aus dem bezüglich der Ordnung letzten Tupel $(Minute_{last}, Attribute_{last})$:

  $$\begin{array}{rcl} Value & := & Attribute_{last} \\ ValueFrom & := & Minute_{last} \end{array}$$

- all_values_within_bounds(+Pat, +Since, +Until, +Attribute, +LowerB, +UpperB)

  Bei Anfragen dieses Typs gibt es keinen Rückgabewert, denn alle Variablen sind beim Aufruf bereits gebunden. Die Anfrage stellt eine zu überprüfende Aussage dar. Der Aufruf schlägt fehl, genau dann, wenn die Aussage falsch ist.

  Die Aussage lautet, dass für den Patienten im angegebenen zeitlichen Intervall das Attribut nur Werte aus dem angegebenen Werteintervall annimmt. Intervallgrenzen für Werte können durch die stellvertretende Angabe der Konstante none ausgelassen werden. Auf diese Weise lassen sich auch einseitig unbeschränkte Intervalle ausdrücken.

  Die Anfrage wird anhand einer SQL-Anfrage der nachfolgenden Form ausgewertet.

[20] Falls die zeitliche Intervallgrenze nicht gesetzt ist, so entfällt die Bedingung „AND (minute > ...)".

```
SELECT COUNT(*) FROM data
WHERE (id=Pat)
AND (minute ≤ Until)
AND (minute > Since)
AND ((Attribute < LowerB) OR (Attribute > UpperB));
```

Bezeichne c den resultierenden Rückgabewert. Dann ist die auszuwertende Aussage genau dann wahr, wenn $c = 0$.

Diese Art von Anfragen wird beispielsweise verwendet, um zu testen, ob in einem mehrstündigen Zeitintervall eine erhobene Patientenvariable durchgängig im Normalintervall liegt.

- any_value_within_bounds(+Pat, +Since, +Until, +Attribute, +LowerB, +UpperB)

  Während das externe Prädikat all_values_within_bounds/6 eine All-Aussage testet, wird bei diesem Prädikat eine analoge Existenz-Aussage über der Datenbank ausgewertet. Anfragen schlagen fehl, genau dann, wenn für das angegebene Zeitintervall alle Werte des Attributs außerhalb des spezifizierten Werteintervalls liegen.

  Die passende SQL-Anfrage lautet:

  ```
  SELECT COUNT(*) FROM data
  WHERE (id=Pat)
  AND (minute ≤ Until)
  AND (minute > Since)
  AND ((Attribute >= LowerB) AND (Attribute <= UpperB));
  ```

  Bezeichne auch hier c den resultierenden Rückgabewert. Die auszuwertende Aussage ist genau dann wahr, wenn $c > 0$.

  Diese Art von Anfragen dient beispielsweise dazu, in zeitlichen Intervallen die Gabe von Medikamenten in bestimmten Dosierungen festzustellen.

- ask_latest_out_of_bounds
  (+Patient, +Since, +Until, +Attribute, +LowerB, +UpperB, -ValFrom, -Val)

  Diese Anfrage entspricht weitestgehend all_values_within_bounds/6, jedoch werden darüber hinaus, falls im Zeitintervall Werte nicht im angegebenen Werteintervall liegen, der letzte solche Zeitpunkt (ValFrom) und der dazugehörige Wert (Val) zurückgegeben.

- sum_of_interval(+Patient, +Since, +Until, +Attribute, -Sum)

  Um beispielsweise die in einem Zeitintervall verabreichte Gesamtmenge einer Infusion zu bestimmen, ist es nötig, alle entsprechenden Einträge im Datensatz aufzusummieren. Die hier beschriebene Anfrage übernimmt diese Aufgabe, wobei neben dem Patientenbezeichner noch das Attribut und das Zeitintervall als Parameter erwartet werden.

  Das Ergebnis wird durch folgende Anfrage bestimmt:

  ```
  SELECT SUM(Attribute) FROM data
  WHERE (id = Patient)
  AND (minute ≤ Until)
  AND (minute > Since);
  ```

- sum_exceeds_during(+Pat, +Since, +Until, +MaxTimeDiff, +MaxSum, +Attribute)

  Diese Anfrage ist weniger elementar, allerdings lässt sie sich nur schlecht unmittelbar in MOBAL realisieren. Sie entspringt in einer weniger parametrisierten Fassung dem Protokoll:

  > „Wurden beim Patienten in den letzten 24 Stunden einmal innerhalb von höchstens 6 Stunden mehr als 6 Liter Infusion verabreicht?"

  Es geht also wieder um die Summe der Einträge eines Attributs für ein Zeitintervall, allerdings werden hier zwei ineinander geschachtelte Zeitfenster betrachtet. Die Anfrage hat die Form einer Aussage, die überprüft wird. Der Aufruf schlägt fehl, genau dann, wenn die Aussage falsch ist. Eine Definition in SQL-Form wird bereits etwas unübersichtlich, daher folgt eine prädikatenlogische Definition. Dabei bezeichne sum_of_interval/5 das oben beschriebene Prädikat der externen Schnittstelle:

  sum_exceeds_during(Pat, Since, Until, MaxTimeDiff, MaxSum, Attribute) $\Leftrightarrow$:

  $$\begin{array}{ll} (\exists t_1)(\exists t_2): & (Since < t_1 < t_2 < Until) \\ \wedge & (t_2 - t_1 < MaxTimeDiff) \\ \wedge & sum\_of\_interval(Patient, t_1, t_2, Attribute, Sum) \\ \wedge & (Sum > MaxSum) \end{array}$$

- ask_latest_change (+Patient, +Timestamp, +Drug, (+/-)in/de-crease, -When, -Diff)

  Diese Anfrage dient speziell der Ermittlung der letzten Veränderung einer Medikamentendosis. Je nachdem, ob increase oder decrease angegeben werden oder die Variable bei der Anfrage ungebunden ist, bezieht sich die Anfrage auf die letzte Dosiserhöhung, Dosisverminderung oder die letzte Änderung der Dosis überhaupt. Es kann also auch gezielt nach der letzten Verminderung gefragt werden, wenn die Dosis danach bereits wieder erhöht wurde.

  Gesucht werden zunächst Zeitpunkte t_1 und t_2 mit den folgenden Eigenschaften:

  $$\begin{array}{ll} & (t_1 < t_2 \leq Timestamp) \\ \wedge & ask\_value(Patient, t_1, Attribute, Value\_Before) \\ \wedge & ask\_value(Patient, t_2, Attribute, Actual\_Value) \\ \wedge & (Value\_Before \neq Actual\_Value) \\ \wedge & (\nexists t_3)(\exists x_3): [\ t_1 < t_3 < t_2 \wedge ask\_value(Patient, t_3, Attribute, x_3)\ ] \\ \wedge & (\forall t_4)(\forall x_4): [\ t_2 < t_4 \leq Timestamp \wedge ask\_value(Patient, t_4, Attribute, x_4) \\ & \rightarrow (x_4 = Actual\_Value)\ ] \end{array}$$

  Es bezeichnet also t_2 den Zeitpunkt der letzten Werteänderung, t_1 den Zeitpunkt des letzten Eintrags davor. Falls ein solches Zeitpunktpaar existiert, so ist es eindeutig und definiert darüber hinaus (Anfrage über ask_value/4) eindeutig die Variablen Actual_Value und Value_Before, wie man sich leicht klar macht. Die Differenz dieser Werte zeigt die Richtung der Änderung an. Ein negativer Wert entspricht der Konstanten decrease, ein positiver der Konstanten increase. War die vierte Komponente der Anfrage ungebunden, oder war sie gebunden und die Änderungsrichtung stimmt überein, so lautet das Ergebnis der Anfrage

  $$\begin{array}{rcl} Diff & := & Actual\_Value - Value\_Before \\ When & := & t_2. \end{array}$$

Andernfalls wird rekursiv eine neue Anfrage mit Timestamp := t_1 durchgeführt. Existieren keine passenden Zeitpunkte t_1 und t_2, so ist der Anfang des Datensatzes erreicht und der Aufruf schlägt fehl.

- ask_oscillating(+Pat, +Time, +Drug, +InnerTimeWindow, +OuterTimeWindow, -NumberOfOscills, -LastOscillsInc)

  Dieses Prädikat lagert eine in MOBAL nur schlecht zu realisierende Anfrage aus, die im Rahmen der Wartezeitenbestimmung (Abschnitt 3.6.9) benötigt wird.

  Für die Semantik sei zunächst der Begriff einer „oszillierenden Dosis“ definiert:

$$
\begin{aligned}
& oscill(Pat, Drug, t_1, t_2) \Leftrightarrow: \\
& (0 < t_2 - t_1 \leq InnerTimeWindow) \\
\wedge\ & ask\_value(Pat, t_1, Drug, Dose\_Decreased) \\
\wedge\ & ask\_value(Pat, t_2, Drug, Dose\_Increased) \\
\wedge\ & Dose\_Decreased < Dose\_Increased \\
\wedge\ & (\forall t_3)(\forall x_3) : [\ t_1 < t_3 < t_2 \wedge ask\_value(Pat, t_3, Drug, x_3) \\
& \rightarrow (x_3 = Dose\_Decreased)\ ] \\
\wedge\ & (\exists t_0) : [\ t_0 < t_1 \wedge ask\_value(Pat, t_0, Drug, Dose\_Before) \\
& \wedge Dose\_Before > Dose\_Decreased \\
& \wedge (\nexists t_4)(\exists x_4) : (t_0 < t_4 < t_1 \wedge ask\_value(Pat, t_4, Drug, x_4))\ ]
\end{aligned}
$$

  Die Zeitpunkte t_1 und t_2 bezeichnen also unmittelbar aufeinander folgende Dosisänderungen, deren zeitlicher Abstand durch die Variable InnerTimeWindow beschränkt ist. Es muss eine Erhöhung der Dosis auf eine Verminderung folgen.

  Bezeichne Osc(Pat, Time, Drug, InnerTimeWindow, OuterTimeWindow) die Menge aller durch Variablen-Substitutionen aus dem Ausdruck oscill(Pat, Drug, t_1, t_2) ableitbaren wahren Grundfakten, für die gilt:

$$(t_1 > Time - OuterTimeWindow) \ \wedge \ (t_2 \leq Time)$$

  Dann ist das Ergebnis der Anfrage wie folgt definiert[21]:

  NumberOfOscills :=
  |Osc(Pat, Time, Drug, InnerTimeWindow, OuterTimeWindow)|

  LastOscillsInc := $\max\{t|$oscill(Pat, Drug, t', t) $\in$
  Osc(Pat, Time, Drug, InnerTimeWindow, OuterTimeWindow)$\}$

3.6.2. Einlesen der Verlaufsdaten

Die beschriebenen externen Prädikate können vom Modell genutzt werden, um die benötigten Informationen zu erfragen. Während komplexere Anfragen über externe Prädikate nur auf höheren Abstraktionsebenen vorkommen, ist das Ziel der Datenschnittstelle auf der untersten Ebene des Modells zunächst nur die Herleitung aller benötigten Patientenvariablen als Fakten eines Prädikates actual_value. Falls für Variablen keine Werte im gültigen Zeitfenster ermittelt werden können, so sollen einfach die entsprechenden Fakten fehlen. Andernfalls soll neben dem gesuchten Variablenwert auch noch der Zeitpunkt der Erhebung angegeben werden, da dieser vom Zeitpunkt der Anfrage abweichen kann. Es ergibt sich die folgende Signatur für das Prädikat:

[21] Die oben mit „+“ gekennzeichneten Variablen werden als gebunden angenommen.

actual_value/5: <patient>, <time>, <attribute>, <value>, <time>.

Bei jedem Fakt dieses Prädikates legen die ersten beiden Komponenten den Patienten und den Zeitpunkt der Aussage fest. Danach wird das Attribut spezifiziert. In der vierten Komponente steht der zum Zeitpunkt der Aussage aktuellste und noch gültige Wert des Attributs. Der Erhebungszeitpunkt des Wertes steht in der letzten Komponente.
Um die Bezeichner für Variablen nicht jeweils im ganzen Modell an die Bezeichner der verwendeten Datenbanken anpassen zu müssen, nimmt ein Prädikat die Abbildung der intern verwendeten Konstanten auf Datenbank-Attribute vor:

attribute_mapping/2: <database_attribute>, <attribute>.

Ob für die Werte einer Variablen eine maximale Verwendbarkeitsdauer besteht und wie lang gegebenenfalls das entsprechende Zeitintervall ist legt das Prädikat

out_of_time_limit/2: <attribute>, <time_interval>.

fest. Existiert für ein Attribut kein Fakt dieses Prädikats, so wird grundsätzlich der aktuellste Wert verwendet. Dies ist im Falle sich nur langsam ändernder, aber selten erhobener Werte sinnvoll, beispielsweise beim Körpergewicht oder der Körperoberfläche.
Die Strategie, nach der Anfragen über eine Dialogkomponente erfolgen, wird gewöhnlich durch die Problemlösungsmethode festgelegt. Im Rahmen der Diplomarbeit kamen lediglich schwache Problemlösungsmethoden zum Einsatz (vgl. S. 28 ff.). Das verwendete Kontrollwissen ist daher nicht vom übrigen Wissen getrennt. Anfragen über die externe Schnittstelle sind in den Regeln des Modells an die Verarbeitung der erfragten Informationen gebunden. Im Falle von Rückwärtsverkettung werden nur dann Anfragen generiert, wenn die zusätzlichen Informationen zur Untersuchung einer fokussierten Lösung notwendig sind. Sollen Widersprüche und Schwachstellen im Modell entdeckt werden, so ist jedoch die Ableitung der inferentiellen Hülle über Vorwärtsverkettung geeigneter. Um in diesem Fall keine Anfragen über die „Dialogkomponente“ mit ungebundenen Variablen zu erzeugen, dient die erste erforderliche Art von Kontrollwissen der Beschränkung von Inferenzen. Zur Illustration folgen hier für einen Ausschnitt der Regeln zur Ableitung aktueller Messwerte beide Varianten.
Betrachten wir zunächst die Variante ohne zusätzliches Kontrollwissen:

```
    attribute_mapping(DB_Attr, Attribute)
∧   out_of_time_limit(Attribute, Limit)
∧   extern(ask_latest_value(Pat, Time, DB_Attr, Limit, Value, From))
⟶   actual_value(Pat, Time, Attribute, Value, From).
```

Bei Anwendung dieser Regel wird zunächst zu jeder im Modell verwendeten Patientenvariable mittels attribute_mapping/2 das passende Datenbank-Attribute, und mittels out_of_time_limit/2 die maximale zeitliche Verwendbarkeit ermittelt. Schließlich wird die Anfrage über das externe Prädikat ask_latest_value/6 durchgeführt. Bei der Verwendung von Rückwärtsverkettung ist dafür zu sorgen, dass bei jeder Anfrage die Variablen *Pat*, *Time* und *Attribute* bereits gebunden sind. Die folgenden Regeln zeigen eine Einarbeitung von Kontrollwissen für die (schwache) Problemlösungsmethode der Vorwärtsverkettung, die eine Auswahl der zu erfragenden Variablen erlaubt.
Soll für einen Patienten zu einem bestimmten Zeitpunkt der aktuellste Wert einer Variablen erhoben werden, so wird dies durch ein Fakt des Prädikats get_attribute_value/3

angezeigt. Ein solcher Zwischenschritt ist sinnvoll, da Werte nicht nur zu Visitenzeitpunkten ermittelt werden. Für jeden Grund, einen Werte zu erheben, genügt nun die Angabe einer Regel mit entsprechender get_attribute_value/3-Konklusion. Ebenso können die verschiedenen Fälle beim Erheben der Werte, etwa mit oder ohne eine maximale Verwendbarkeit, sich auf die Fakten dieses Prädikats beziehen, was hilft, mehrere Fälle mit den gleichen Regeln abzudecken.
Der obige Ausschnitt mit den passenden Ergänzungen hat die folgende Form[22]:

```
    visit(Pat, Time)
∧   attrib(Attribute)
⟶   get_attribute_value(Pat, Time, Attribute).

    get_attribute_value(Pat, Time, Attribute)
∧   attribute_mapping(DB_Attr, Attribute)
∧   out_of_time_limit(Attribute, MaxDiff)
∧   extern(ask_latest_value(Pat, Time, DB_Attr, MaxDiff, Val, ValFrom))
⟶   actual_value(Pat, Time, Attribute, Val, ValFrom).
```

Einige Variablen werden aus mehreren Patientenvariablen zusammengesetzt. Beispielsweise fehlt der Herzindex (Cardiac Index) in dem Datensatz, der in Abschnitt 4.1 für Experimente verwendet wird. Da der Herzindex als Quotient aus Herzzeitvolumen (Cardiac Output) und Körperoberfläche (Body Surface Area) definiert ist, kann er durch die Angabe folgender Regeln berechnet werden[23]:

```
    get_attribute_value(Patient, Timestamp, cardiac_index)
⟶   get_attribute_value(Patient, Timestamp, cardiac_output).

    get_attribute_value(Patient, Timestamp, cardiac_index)
⟶   get_attribute_value(Patient, Timestamp, body_surface_area).

    actual_value(Pat, Time, cardiac_output, CO, CO_From)
∧   actual_value(Pat, Time, body_surface_area, BSA, BSA_From)
∧   div(CO, BSA, CI)
⟶   actual_value(Patient, Timestamp, cardiac_index, CI, CO_From).
```

3.6.3. Werteänderungen

Zu den nicht zu Visitenzeitpunkten erhobenen Werten zählen besonders die „Referenzwerte“. Diese werden vom Protokoll zu einigen Patientenvariablen erhoben, um deren Veränderung über die Zeit zu bewerten. Der Referenzwert zu einem Messwert ist bei einer Visite der aktuellste, mindestens 15 Minuten alte Wert.
Die Erhebung geschieht über das im letzten Abschnitt vorgestellte get_attribute_value/3. Da die „aktuellen“ Messwerte teilweise schon Stunden alt seien können, bezieht sich in der Wissensbasis die Mindestdifferenz von 15 Minuten auf den *Erhebungs*zeitpunkt des

[22] In MOBAL lassen sich Sorten über unäre Prädikate repräsentieren. Das Prädikat attrib/1 bestimmt hier die Sorte aller zu erfragender Attribute.
[23] Dabei führt div(CO, BSA, CI) die Berechnung $CI := CO/BSA$ durch.

aktuellen Wertes, der jeweils im Prädikat actual_value/5 vermerkt ist. Zunächst ist also für jede Visite und für jedes auf diese Weise bewertete Attribut (*reference_attribute*) der Referenzzeitpunkt zu errechnen. Für die notwendige Subtraktion steht das MOBAL built-in

sub/3: <arg1>, <arg2>, <arg3>.

zur Verfügung, das die Berechnung $Arg_3 := Arg_1 - Arg_2$ durchführt. Für den ermittelten Referenzzeitpunkt wird der aktuellste Wert dann wie für Visitenzeitpunkte erhoben. Die maximale zeitliche Verwendbarkeit von Werten findet auch für Referenzwerte Anwendung[24]. Dies dient dazu, wenig aussagekräftige Vergleiche zu vermeiden, bei denen aktuelle Werte mit viel älteren Werten verglichen werden.
Die Berechnung des Referenzzeitpunktes und die Erhebung des Wertes geschehen in der Wissensbasis anhand der folgenden Regeln[25]:

```
     visit(Patient, Now)
 ∧   reference_attribute(Attribute)
 ∧   actual_value(Patient, Now, Attribute, Value, Value_From)
 ∧   sigchange_min_time(MinDiff)
 ∧   sub(Value_From, MinDiff, Reference_Time)
⟶   reference_value_from(Patient, Now, Attribute, Reference_Time).

     reference_value_from(Patient, Now, Attribute, Reference_Time)
⟶   get_attribute_value(Patient, Reference_Time, Attribute).
```

Falls für ein Attribut sowohl zum Visitenzeitpunkt, als auch zum passenden Referenzzeitpunkt ein Wert bekannt ist, so wird als Grundlage für die weitere Bewertung die absolute Differenz dieser Werte berechnet. Dabei findet keinerlei „Glättung" von Messwerten statt. Selbst eventuell eine Minute vor oder nach den fraglichen Zeitpunkten erhobene Messwerte haben keinerlei Einfluss auf das Ergebnis, nicht einmal wenn sie deutlich von den verwendeten Werten abweichen.
Die Differenz wird für die weitere Verarbeitung als Fakt des Prädikats

absolute_value_diff/4: <patient>, <time>, <reference_attribute>, <val_diff>

abgelegt. Dies geschieht mit Hilfe der nachfolgenden Regel:

```
     reference_value_from(Patient, Now, Attribute, RefTime)
 ∧   actual_value(Patient, Now, Attribute, ValueNow, Now_From)
 ∧   actual_value(Patient, RefTime, Attribute, RefValue, Ref_From)
 ∧   sub(ValueNow, RefValue, Diff)
⟶   absolute_value_diff(Patient, Now, Attribute, Diff).
```

[24] Die Haltbarkeit bezieht sich dann allerdings auf den berechneten Referenzzeitpunkt, nicht auf den Visitenzeitpunkt.

[25] Über das Prädikat sigchange_min_time/1 wird dabei die konstante Zeitdifferenz von 15 Minuten festgelegt.

3.6.4. Qualitative Abstraktionen

Die Daten, die erhobene Messwerte und Medikationen beschreiben, sind üblicherweise numerischer Natur. Die auf Seite 26 erwähnten qualitativen Abstraktionen bilden eine Möglichkeit, aus diesen Merkmalen für den weiteren Inferenzprozess günstigere Merkmalsinterpretationen zu gewinnen. Das Behandlungsprotokoll bildet zunächst intervallbasiert die numerischen Ausgangsdaten auf Kategorien ab. Den Therapie-Entscheidungen liegt eine symbolische Repräsentation des Patientenzustands zugrunde. Qualitative Abstraktionen, bei denen jeder Wert intervallbasiert genau einer Kategorie zugeordnet wird, werden im folgenden als *Diskretisierung* bezeichnet.

Diskretisierungen beziehen sich in der Wissensbasis auf Medikationen, Messwerte und die Änderung von Messwerten im zeitlichen Verlauf. Hinter den Ableitungen stehen teilweise verschiedene Regeln, was jedoch eher technische Gründe hat. So sind die Intervalle teils halboffen, teils geschlossen und mitunter dürfen Intervallgrenzen fehlen. Die Konstanten für die Kategorien und die Prädikats-Bezeichner variieren. Mal genügt die Angabe eines Normalintervalls, mal ist eine feinere disjunkte Zerlegung mit entsprechend mehr Kategorien notwendig. Trotz dieser Unterschiede ist deutlich die gemeinsame Methode zu erkennen.

Zunächst sei die Diskretisierung der Patientenvariablen erwähnt. Die zu diskretisierende Teilmenge dieser Werte wird über discr_attrib/1 festgelegt. Die Ergebnisse werden als Fakten des Prädikats

discrete_value/4: <patient>, <time>, <attribute>, <category>.

repräsentiert.

Die folgenden Regeln[26] geben die Rahmenbedingungen für Diskretisierungen laut Definition wieder[27]:

visit(Patient, Timestamp)
$\wedge$ discr_attrib(Attribute)
$\leadsto$ discrete_value(Patient, Timestamp, Attribute, Category).

discrete_value(Patient, Timestamp, Attribute, Category1)
$\wedge$ discrete_value(Patient, Timestamp, Attribute, Category2)
$\leadsto$ Category1 = Category2.

visit(Patient, Timestamp)
$\wedge$ discr_attrib(Attribute)
$\wedge$ \+(actual_value(Patient, Timestamp, Attribute, Value, Value_From))
$\longrightarrow$ discrete_value(Patient, Timestamp, Attribute, unknown).

Die MOBAL-Constraints stellen sicher, dass zu jeder Visite und jeder zu diskretisierenden Variable genau ein Fakt existiert, das dem Wert eine Kategorie zuordnet. Die Constraints dienen ausschließlich dem Aufspüren von Fehlern im Modell. Wenn etwa mehrere Bewertungen für eine Variable abgeleitet werden, so setzt MOBAL die entsprechenden Fakten

[26] Hier wird der Begriff „Regel“ auch für Constraints gebraucht. Letztere sind an dem Ableitungssymbol „$\leadsto$“ zu erkennen.

[27] Der Ausdruck „A = B“ steht für „eq(A, B)“ in MOBAL, also für „A und B unifizierbar“.

auf eine sogenannte „Agenda“. Gefundene Fehler können zu einem passenden Zeitpunkt unter Zuhilfenahme der MOBAL-Werkzeuge korrigiert werden.
Ist zu einer zu diskretisierenden Variable ein Wert nicht verfügbar, abzulesen am fehlenden actual_value/5-Fakt, so wird die ausgezeichnete Kategorie unknown gewählt. Der Vorteil einer eigenen Kategorie für diesen Fall liegt darin, dass dies hilft, die Anzahl nicht-monotoner Schlüsse in den weiteren Regeln zu minimieren. Auf diesen Aspekt wird in Abschnitt 3.8.2 noch ausführlicher eingegangen.
Die Diskretisierung von Patientenvariablen lässt sich unterteilen in solche, die auf einem Normalintervall basieren, und solche, die Messwerte in mehr als drei Kategorien unterteilen. Normalintervalle stellen eine kompakte Art dar, Wissen standardisiert auszudrücken. So sagt das Fakt

```
normal_interval(central_venous_pressure, 5.0, 14.0)
```

beispielsweise aus, dass der zentralvenöse Blutdruck im Intervall [5, 14] liegen sollte[28]. Es lassen sich nun standardisierte Konstanten für die resultierenden Kategorien angeben. Ist der Wert zu niedrig, so wird die Kategorie low gewählt. Zu hohe Werte werden auf high abgebildet. Werte innerhalb des Intervalls gehören zur Kategorie normal. Um die Flexibilität zu erhöhen, dürfen auch Intervallgrenzen fehlen. Hierfür dient die Konstante none.
Zur Diskretisierung mit mehr als drei Kategorien, oder falls die im Rahmen der Normalintervalle verwendeten Bezeichner unzutreffend sind, dient ein weiterer, weniger restriktiver Ansatz. Als Beispiel soll die Diskretisierung des mittleren arteriellen Blutdrucks dienen:

```
discrete_interval(mean_arterial_pressure,  none,   60.0,  threatening_low_bp).
discrete_interval(mean_arterial_pressure,  60.0,   65.0,  very_low_bpav).
discrete_interval(mean_arterial_pressure,  65.0,   70.0,  low_bp).
discrete_interval(mean_arterial_pressure,  70.0,   90.0,  adequate_bp).
discrete_interval(mean_arterial_pressure,  90.0,  120.0,  normal_bp).
discrete_interval(mean_arterial_pressure, 120.0,   none,  excessive_bp).
```

Bei diesen Fakten stehen die Intervallgrenzen an der zweiten und dritten Stelle. Die Kategorie lässt sich nun nicht mehr, wie bei Normalintervallen, automatisch ableiten, sondern steht jeweils an vierter Stelle der Fakten. Um eine disjunkte Zerlegung zu erhalten, müssen die Intervalle als halboffen interpretiert werden.
Die Interpretation des Fakts zum „normalen Blutdruck“ ist nun beispielsweise wie folgt:

> Falls der arterielle Mitteldruck im Intervall [90, 120) liegt, so gehört er zur Kategorie normal_bp.

Die Konstante none steht wieder für eine fehlende Intervallgrenze. So wird die Kategorie threatening_low_bp bei allen Werten kleiner 60 zugeordnet.
Während die Fakten das Sachbereichswissen angeben kann man die Regeln, die mit Hilfe dieser Fakten die numerischen Eingabewerten auf diskrete Kategorien ableiten als Kontrollwissen auffassen (vgl. Abschnitt 3.3.2). Die bereits erwähnten MOBAL-Constraints und die Regel für unbekannte Werte stellen einen Teil dieses Kontrollwissens dar. Darüber hinaus sind Regeln notwendig, die Vergleiche anstellen.

[28] Die Einheit ist hier mmHg. Für jede einzelne Variable wird nur eine Einheit verwendet, so dass diese nicht gesondert vermerkt werden muss.

Die Regeln für Normalintervalle, bei denen beide Intervallgrenzen existieren, sehen wie folgt aus[29]:

actual_value(Patient, Timestamp, Attribute, Value, From)
∧ normal_interval(Attribute, LowerBound, UpperBound)
∧ Value < LowerBound
⟶ discrete_value(Patient, Timestamp, Attribute, low).

actual_value(Patient, Timestamp, Attribute, Value, From)
∧ normal_interval(Attribute, LowerBound, UpperBound)
∧ Value > UpperBound
⟶ discrete_value(Patient, Timestamp, Attribute, high).

actual_value(Patient, Timestamp, Attribute, Value, From)
∧ normal_interval(Attribute, LowerBound, UpperBound)
∧ Value ≥ LowerBound)
∧ Value ≤ UpperBound
⟶ discrete_value(Patient, Timestamp, Attribute, normal).

Der analoge Fall für die Diskretisierung anhand von discrete_interval/4-Fakten:

actual_value(Patient, Timestamp, Attribute, Value, Value_From)
∧ discrete_interval(Attribute, LowerBound, UpperBound, Category)
∧ Value ≥ LowerBound
∧ Value < UpperBound
⟶ discrete_value(Patient, Timestamp, Attribute, Category).

Die Diskretisierung der Medikamentendosierungen ähnelt der Diskretisierung über Normalintervalle. Es gibt für jedes Medikament einen normalen Dosisbereich. Nach unten ist dieser Bereich durch die Grenze definiert, ab der das Medikament eine Wirkung zeigt. Eine Überschreitung der Obergrenze stellt eine Überdosierung dar. Das Wissen über diese Intervalle ist in der Form von Fakten des Prädikats

drug_dose_bounds/3: <drug>, <dose>, <dose>.

abgelegt. Die beiden letzten Stellen geben dabei die untere und obere Schranke für die Dosis des angegebenen Medikaments an. Die Diskretisierung unterscheidet sich von der über Normalintervalle besonders durch die unterschiedlichen Kategorien. Eine Dosis von Null wird auf die Kategorie none abgebildet. Unterhalb der unteren Schranke heißt die Kategorie under_dose, oberhalb der oberen over_dose. Innerhalb des geschlossenen Intervalls wird die Dosis als normal_dose bezeichnet. Wie auch bereits bei den anderen Diskretisierungen existiert für unbekannte Werte eine eigene Kategorie namens unknown. Die Regeln, die die drug_dose_bounds/3-Fakten auswerten, unterscheiden sich nur geringfügig von den bereits vorgestellten Regeln. Das Zielprädikat ist allerdings

discrete_drug_dose/4: <patient>, <time>, <drug>, <discr_dose>.

[29] Dabei sind *ge gt*, *le* und *lt* MOBAL built-ins für Vergleiche. ge(A, B) steht für A ≥ B (greater or equal). Analog steht gt für „>“, lt für „<“ und le für „≤“

Dies dient besonders der Übersichtlichkeit der erzeugten Faktenmenge zu jeder Visite, da die Medikation so besonders leicht zu erkennen ist[30].
Eine weitere Diskretisierung bezieht sich auf die im zeitlichen Verlauf auftretende Veränderung von Messwerten. Wie in Abschnitt 3.6.3 bereits beschrieben wurde, wird für jede Visite und jede reference_attribute-Konstante die Differenz des aktuellen Wertes zu einem älteren Wert berechnet. Diese wird jeweils als Fakt des Prädikats absolute_value_diff/4 abgelegt. Auf der hier beschriebenen Ebene wird eine Bewertung der Differenz vorgenommen, die als Fakt des Prädikats

significant_change/4: <patient>, <time>, <attribute>, <change>.

repräsentiert wird. Dabei bilden die Konstanten der Sorte <change> die möglichen Bewertungen. Außer einem signifikanten Anstieg (Konstante: increase) kann noch eine signifikante Verminderung (decrease) und eine nicht signifikante Veränderung (stable) hergeleitet werden. Falls keine absolute Differenz ableitbar ist, weil für den Visitenzeitpunkt oder den Referenzzeitpunkt kein Wert verfügbar ist, so wird die Konstante unknown gewählt. Dies entspricht also ganz den bisher vorgestellten Diskretisierungen.
Eine Änderung gilt als signifikant, wenn der Betrag der Differenz einen bestimmten Wert erreicht. Dieser Mindestbetrag hängt von der betrachteten Variablen ab und wird anhand des Prädikats

significant_change_diff/2: <attribute>, <val_diff>.

festgelegt. Exemplarisch sei die Ableitung eines signifikanten Anstiegs angeführt:

absolute_value_diff(Patient, Timestamp, Attribute, Diff)
$\wedge$ significant_change_diff(Attribute, MinDiff)
$\wedge$ Diff $\geq$ MinDiff
$\longrightarrow$ significant_change(Patient, Timestamp, Attribute, increase).

Die anderen Bewertungen werden analog abgeleitet. Die Regeln, die sicherstellen, dass jeder Wert genau einer Kategorie zugeordnet wird sehen wie folgt aus:

visit(Patient, Timestamp)
$\wedge$ reference_attribute(Attribute)
$\leadsto$ significant_change(Patient, Timestamp, Attribute, Change).

significant_change(Patient, Timestamp, Attribute, Change1)
$\wedge$ significant_change(Patient, Timestamp, Attribute, Change2)
$\leadsto$ Change1 = Change2.

visit(Patient, Timestamp)
$\wedge$ reference_attribute(Attribute)
$\wedge$ \+(absolute_value_diff(Patient, Timestamp, Attribute, Diff))
$\longrightarrow$ significant_change(Patient, Timestamp, Attribute, unknown).

[30] Da für jede Visite recht viele Fakten abgeleitet werden, ist auch dem Aspekt der Übersichtlichkeit Rechnung zu tragen. Dies gilt besonders, wenn bei eventuell zeitlich knapp bemessenen Treffen mit einem Experten Fälle diskutiert werden sollen. In MOBAL lassen sich Fakten sortiert ausgeben, so dass alle Fakten zum gleichen Prädikat aufeinander folgen.

Schließlich kommt im Modell noch eine Form von qualitativer Abstraktion vor, die gemäß der verwendeten Definition nicht unter den Begriff der Diskretisierung fällt. Medikamente sollen, in Abhängigkeit von der Dosis, Wirkungsgruppen zugeordnet werden. Es werden zwar intervallbasiert für numerische Werte Kategorien hergeleitet, jedoch können dabei auch mehrere, oder gar keine Kategorien abgeleitet werden. Grundlage der Zuordnung ist die Tabelle 3.1 auf Seite 19, die in der Wissensbasis über das Prädikat

drugdose_to_category/4: <drug>, <dose>, <dose>, <drug_category>.

repräsentiert wird. Die Dosierungen an zweiter und dritter Stelle bilden ein geschlossenes Intervall innerhalb dessen das Medikament an erster Stelle in die Wirkungsgruppe an vierter Stelle fällt. Die Konstante none darf wieder verwendet werden um anzuzeigen, dass das Intervall unbeschränkt ist. Fehlt die Untergrenze, so wird „echt größer Null“ als Bedingung angenommen.
Als Beispiel sei Adrenalin[31] angeführt:

```
drugdose_to_category(supra,   none,   0.15,   inotrope).
drugdose_to_category(supra,   0.03,   none,   pressor).
```

Wie im Abschnitt 3.2.2 beschrieben, ändert sich die Wirkung von Adrenalin dosisabhängig. Im Bereich von 0.03 bis 0.15 μg/kg/min ist die Wirkung sowohl positiv inotrop, als auch gefäßverengend. Eine disjunkte Zerlegung des Dosisbereichs, um jeweils genau eine Wirkungsgruppe zu erhalten, ist demnach nicht möglich. Im Gegensatz dazu ist Dopamin in einer Dosierung unterhalb von 3 μg/kg/min keiner im Protokoll benannten Gruppe zugehörig, wird aber in Deutschland durchaus in dieser Dosierung verwendet um die Nierenfunktion zu verbessern. Für die Modellierung wurde entsprechend auf die Regeln sowohl zur Existenz, als auch zur Eindeutigkeit einer Kategorie verzichtet, was den Unterschied zur Diskretisierung auch formal unterstreicht.
Die Ergebnisse der Auswertung werden in Fakten des Prädikats

drug_effect_category/4: <patient>, <time>, <drug_category>, <drug>.

abgelegt. Die Regeln zur Herleitung entsprechen weitestgehend denen zur Diskretisierung. Es sei noch angemerkt, dass die Zuordnungen von Dosisbereichen zu Wirkungen hier eher im Sinne von Definitionen zu verstehen sind. Dass tatsächlich immer die erwartete Wirkung auf die entsprechenden Parameter eintritt, ist keinesfalls die Aussage dieser Zuordnungen.

[31] Als Medikament „Suprarenin“, daher die Konstante supra.

3.6.5. Bewertung von Patientenzustand und Medikation

Im letzten Abschnitt wurden qualitative Abstraktionen von Variablen als eine Möglichkeit beschrieben, für weitere Inferenzen kompaktere Beschreibungen zu gewinnen. Darüber hinaus finden sich im Behandlungsprotokoll weitere Abstraktionsmechanismen. Hierzu zählen insbesondere Definitionen im Sinne von Seite 25, die einer Bewertung des Patientenzustands dienen. Die Eingabe für diese Schicht stellen in erster Linie bereits diskretisierte Variablen dar. Darüber hinaus werden für einige Begriffe Informationen direkt über die Datenschnittstelle bezogen. Dies bezieht sich beispielsweise auf Anfragen der Art

> „Lag der mittlere arterielle Blutdruck des Patienten innerhalb der letzten 24 Stunden einmal unter dem Wert von 70 mmHg?“

Für diese Anfrage müssen alle Datenbanktupel des Patienten ausgewertet werden, deren Attribut minute im untersuchten Zeitintervall liegt. Die Daten der bereits in der Vergangenheit ausgewerteten Visitenzeitpunkte reichen nicht aus. Hinzu kommt, dass in MOBAL Regeln wie die folgende unvertretbare Performanzprobleme mit sich bringen[32]:

$$\begin{array}{ll}
 & \mathsf{visit(Patient,\ Timestamp)} \\
\wedge & \mathsf{actual_value(Patient,\ When,\ mean_arterial_pressure,\ MAP,\ From)} \\
\wedge & \mathsf{MAP} < \mathsf{70} \\
\wedge & \mathsf{sub(Timestamp,\ 1440,\ Yesterday)} \\
\wedge & \mathsf{When} \leq \mathsf{Timestamp} \\
\wedge & \mathsf{When} > \mathsf{Yesterday} \\
\longrightarrow & \mathsf{recently_hypotension(Patient,\ Timestamp).}
\end{array}$$

Das Kritische an dieser Regel ist, dass die Menge zunächst betrachteter Fakten des Prädikats actual_value/5 nicht hinreichend beschränkt ist, da die Variable When bei der Auswertung ungebunden ist. Effizienter ist bei solchen Anfragen die Verwendung eines geeigneten Prädikats der externen Datenschnittstelle.

Das Modell verwendet eine Reihe von Merkmalen, die sich aus einer kontinuierlichen Auswertung meist mehrstündiger Intervalle über der Datenbank ergeben. Es handelt sich dabei um Kriterien, anhand derer das Protokoll die Stabilität des Patientenzustands bewertet.

Tabelle 3.3 gibt eine Übersicht über die überprüften Ereignisse und die zugeordneten Terme der Sorte <unstability_term>. Aus Gründen der Übersichtlichkeit werden diese Merkmale als Fakten eines gemeinsamen Prädikats

> unstability_criteria/3: <patient>, <time>, <unstability_term>

repräsentiert.

Das Behandlungsprotokoll unterscheidet an dieser Stelle nicht zwischen fehlenden Werten und der Negation des Ereignisses, sondern nur zwischen „Ereignis ableitbar“ und „Ereignis nicht ableitbar“. Zur Vereinfachung werden auch im umgesetzten Modell daher lediglich positive Literale des Prädikats unstability_criteria/3 abgeleitet. Ob ein Ereignis *nicht* eingetreten ist kann mit Hilfe des Operators „\+“ abgefragt werden. Bei Verwendung dieses Operators können Ableitungen zwar ungültig werden, wenn neue Fakten hinzugefügt werden, jedoch hängt die Ableitung der hier betrachteten Fakten nur von den als

[32] Zeitintervalle werden in Minuten beschrieben. 1440 ist die Zahl der Minuten pro Tag.

Konstante	Ereignis
recently_hypotension	Der mittlere arterielle Blutdruck lag innerhalb von 24 Stunden einmal unter 70 mmHg.
recently_much_infu	Innerhalb von 24 Stunden wurden mindestens 6 Liter Infusion in höchstens 6 Stunden gegeben.
recently_pressor	Innerhalb der letzten 24 Stunden wurde ein Vasokonstriktor gegeben.
recently_inotrope	Innerhalb der letzten 24 Stunden wurde ein positiv inotropes Medikament verabreicht.
recently_nsbolus	Innerhalb der letzten 12 Stunden lief eine Infusion als Bolus. Heuristik zum Bolus: Summe der Infusionen einer Stunde überstieg 15 ml/kg Körpergewicht.
recently_lasix	Lasixgabe vor höchstens 1 Stunde, in Bolusform vor höchstens 6 Stunden
recently_diuretics	Lasixgabe vor höchstens 6 Stunden, egal ob per Perfusor oder als Bolus
recently_excessive_glucose	Der letzte 15 Minuten bis 24 Stunden alte Blutzuckerwert ist größer als 300 mg/dl.
recently_mannitol	Gabe von Mannitol innerhalb der letzten 2 Stunden

Tabelle 3.3.: Anhand der Datenschnittstelle überprüfte Ereignisse zur Patientenstabilität

konstant angenommenen Daten aus der Datenbank ab. Ableitungsgraphen in MOBAL werden durch die Verwendung des Operators „\+“ unvollständig. In diesem Fall fehlt allerdings nur die Anfrageregel. Diese kann leicht über die Prämisse in der Regelmenge identifiziert werden. Das Thema der autoepistemischen Operatoren und nicht-monotonen Schlüsse wird in Abschnitt 3.8.2 noch ausführlicher diskutiert.

Die folgenden Regeln zeigen exemplarisch, wie sich die Ereignisse zu den Konstanten recently_inotrope und recently_much_infu mit Hilfe der in Abschnitt 3.6.1 definierten externen Prädikate überprüfen lassen. Die Regeln sind dabei gegenüber dem Modell vereinfacht, da nur die Idee skizziert werden soll. Die Werte 360 und 1440 geben Zeitspannen in Minuten an, die 6000 entsprechen den 6 Litern Infusionsvolumen.

```
    visit(Pat, Now)
∧   drugdose_to_category(Drug, LowB, UpB, pressor)
∧   sub(Now, 1440, Start)
∧   extern(any_value_within_bounds(Pat, Start, Now, Drug, LowB, UpB))
⟶   unstability_criteria(Pat, Now, recently_pressor).

    visit(Pat, Now)
∧   sub(Now, 1440, Start)
∧   extern(sum_exceeds_during(Pat, Start, Now, 360, 6000, total_input))
⟶   unstability_criteria(Pat, Now, recently_much_infu).
```

Auf ähnliche Weise erhoben werden die jeweils letzten Medikamentendosierungen größer

Null innerhalb eines bestimmten Zeitfensters, die Zeitpunkte der letzten Dosisänderungen und deren Richtung. Die zugehörigen Prädikate sind

last_known_dose/5: <patient>, <time>, <drug>, <dose>, <time>.

last_dose_change/5: <patient>, <time>, <drug>, <change>, <time>.

Die jeweils ersten Zeitpunkte beziehen sich auf den Moment der Auswertung (Visite), die jeweils letzten Stellen geben hingegen den Zeitpunkt des erfragten Ereignisses an.
Neben diesen Abfragen finden sich Abstraktionen bereits erhobener Variablen im Modell. Dabei wird hauptsächlich aus mehreren, bereits diskretisierten Werten eine umfassendere Bewertung abgeleitet. Hierfür sei die arterielle Durchblutung („arterial circulation") als Beispiel angeführt. Die Bewertung dieser Größe und ihrer Veränderung bezieht sich nicht auf einen speziellen Messwert, sondern es kommen mehrere Attribute als Indikatoren in Frage. Die Daten eines PA-Katheters sind zuverlässiger, jedoch seltener verfügbar als klinische Variablen.
In der vorliegenden Version des Protokolls wird bereits bei einem Indikator für unzureichende arterielle Durchblutung die arterielle Durchblutung insgesamt als unzureichend eingestuft. Ist keiner der Indikatoren ableitbar, so lautet die Bewertung „Durchblutung hinreichend". Die Bewertung entspricht also einer einfachen disjunktiven Verknüpfung. Bei der im Anschluss vorgestellten Bewertung der *Änderung* der Durchblutung wird hingegen die unterschiedliche Zuverlässigkeit und Aussagekraft der Daten in Form einer Hierarchie berücksichtigt.
Die zur Bewertung herangezogenen Indikatoren werden in der Wissensbasis anhand des Prädikats

ineffective_art_circ_evidence/2: <attribute>, <category>.

repräsentiert und sind in Tabelle 3.4 zusammengefasst. Das erste Argument gibt jeweils einen Messwert an, das zweite die diskrete Kategorie in die dieser fallen muss.
Das Kontrollwissen realisiert wie folgt die disjunktive Verknüpfung der als Fakten repräsentierten Indikatoren:

```
    visit(Patient, Timestamp)
∧   ineffective_art_circ_evidence(Attribute, Category)
∧   discrete_value(Patient, Timestamp, Attribute, Category)
⟶   ineffective_art_circ_data(Patient, Timestamp, Attribute).

    ineffective_art_circ_data(Patient, Timestamp, Attribute)
⟶   ¬effective_arterial_circulation(Patient, Timestamp).

    visit(Patient, Timestamp)
∧   \+(ineffective_art_circ_data(Patient, Timestamp, Attribute))
⟶   effective_arterial_circulation(Patient, Timestamp).
```

Dass die abgeleiteten Indikatoren als Konstanten in einem „Zwischenprädikat" abgelegt werden ermöglicht eine leichte Anpassung an komplexere Kompositionen der Zwischenergebnisse bei Ergänzungen der Wissensbasis.
Zu den komplexeren Auswertungsmethoden des Protokolls zählt die wiederholt auftretende Methode, aus einer hierarchisch geordneten Menge das erste verfügbare bzw. zutreffende Element zu wählen. Dies soll nun am Beispiel der Änderung der arteriellen

Klinische Variablen:		
ineffective_art_circ_evidence(	lactate,	high).
ineffective_art_circ_evidence(	urine_output,	low).
ineffective_art_circ_evidence(	osmolality,	high).
ineffective_art_circ_evidence(	extremities,	bad_state).
ineffective_art_circ_evidence(	heartrate,	hr_high).
ineffective_art_circ_evidence(	heartrate,	hr_very_high).
PA-Katheter-Variablen:		
ineffective_art_circ_evidence(	cardiac_index,	low).
ineffective_art_circ_evidence(	svo2,	low).
ineffective_art_circ_evidence(	pvo2,	low).
ineffective_art_circ_evidence(	cavo2,	high).

Tabelle 3.4.: Indikatoren für eine unzureichende arterielle Durchblutung

Durchblutung im zeitlichen Verlauf erläutert werden. Grundlage für die Bewertung sind diskretisierte Änderungen (vgl. S. 45) anderer Patientenvariablen.
Die Hierarchie der Variablen und die jeweilige Quelle[33] wird über das Prädikat

arterial_circulation_hierarchy/3: <attribute>, <source>, <art_circ_id>.

repräsentiert. Die zugewiesenen Werte der Sorte <art_circ_id> fungieren sowohl als Prioritäten, die eine Ordnung über den Attributen definieren, als auch als IDs, anhand derer die Attribute eindeutig identifiziert werden können. Tabelle 3.5 gibt die Hierarchie wieder.
Die zugehörigen Regeln[34] dienen dazu, die nach der Tabelle hierarchiehöchste verfügbare Bewertung als Gesamtbewertung abzuleiten. Die Eindeutigkeit wird mit Hilfe eines Constraints sichergestellt:

$$\begin{array}{ll} & \mathsf{arterial\_circulation\_hierarchy(Attribute1,\ Source1,\ ID)} \\ \wedge & \mathsf{arterial\_circulation\_hierarchy(Attribute2,\ Source2,\ ID)} \\ \rightsquigarrow & \mathsf{Attribute1 = Attribute2.} \end{array}$$

$$\begin{array}{ll} & \mathsf{visit(Pat,\ Time)} \\ \wedge & \mathsf{arterial\_circulation\_hierarchy(Attribute,\ Source,\ ID)} \\ \wedge & \mathsf{significant\_change(Pat,\ Time,\ Attribute,\ Direction)} \\ \wedge & \mathsf{Direction} \neq \mathsf{unknown} \\ \longrightarrow & \mathsf{arterial\_circ\_change\_evidence(Pat,\ Time,\ Attribute,\ ID,\ Direction).} \end{array}$$

[33] Die Aufgabe der im Protokoll vorgenommenen Unterscheidung nach Datenquellen wird im hier beschriebenen Kontext bereits durch die Hierarchiewerte erfüllt. Die zusätzliche Information der Quelle erleichtert jedoch Änderungen.

[34] Der besseren Übersicht halber wurden in der dritten Regel nur einmal verwendete Variablen durch einen Unterstrich ersetzt.

	Variable	*Datenquelle*	*Priorität*
arterial_circulation_hierarchy(	cardiac_index,	pa_data,	7).
arterial_circulation_hierarchy(	pvo2,	pa_data,	6).
arterial_circulation_hierarchy(	svo2,	pa_data,	5).
arterial_circulation_hierarchy(	cavo2,	pa_data,	4).
arterial_circulation_hierarchy(	lactate,	clinical,	3).
arterial_circulation_hierarchy(	extremities,	clinical,	2).
arterial_circulation_hierarchy(	heartrate,	clinical,	1).

Tabelle 3.5.: Variablen-Hierarchie zur Bewertung der Entwicklung der arteriellen Durchblutung.

```
    arterial_circ_change_evidence(Pat, Time, _, _, _)
∧   max_of(I, arterial_circ_change_evidence(Pat, Time, _, I, _), ID)
∧   arterial_circ_change_evidence(Pat, Time, _, ID, Direction)
⟶   arterial_circulation_change(Pat, Time, Direction).

    visit(Patient, Time)
∧   \+(arterial_circ_change_evidence(Patient, Time, Attr, ID, Dir))
⟶   arterial_circulation_change(Patient, Time, unknown).
```

Zu allen in der Hierarchie vorkommenden Variablen werden zunächst verfügbare signifikante Veränderungen im Prädikat arterial_circ_change_evidence/5 „eingesammelt". Mit Hilfe des autoepistemischen Operators max_of[35] wird dann der Eintrag höchster Priorität ermittelt. Falls für keine der Variablen eine Bewertung der Veränderung verfügbar ist, so weist die letzte Regel die Kategorie unknown zu.
Die möglichen Bewertungen im Zielprädikat

arterial_circulation_change/3: <patient>, <time>, <change>.

entsprechen der Sorte <change>, die bereits vom Prädikat significant_change/4 (S. 45) bekannt ist. Bei jeder Visite wird genau eine solche Bewertung abgeleitet.
Während sich solche Begriffe noch gut über eine Menge von Fakten darstellen lassen, die dann anhand weniger Regeln interpretiert werden, ist dies beispielsweise beim Begriff urine_output_reliable/2, der die Verlässlichkeit des Urinflusses anhand verschiedener Merkmale bewertet, problematisch. Daher wird dieser Begriff im Modell über eine Menge einfacher Regeln abgeleitet, die jeweils wenige Prämissen besitzen, jedoch nicht unbedingt nur eine, wie dies implizit etwa im Beispiel der effective_arterial_circulation/2 der Fall war.
Ein Teil der zugehörigen Regeln lautet:

```
    unstability_criteria(Patient, Timestamp, recently_diuretics)
⟶   ¬urine_output_reliable(Patient, Timestamp).
```

[35] Der Operator max_of(Var, L(..., Var, ...), Result) ermittelt aus der Menge aller mit *L(..., Var, ...)* unifizierbaren Literale die maximale Belegung der Variablen Var. Result wird mit dem Ergebnis unifiziert.

$$\begin{array}{rl} & \text{discrete\_value(Patient, Timestamp, creatinin, very\_high)} \\ \longrightarrow & \neg\text{urine\_output\_reliable(Patient, Timestamp).} \end{array}$$

$$\begin{array}{rl} & \text{discrete\_value(Patient, Timestamp, creatinin, high)} \\ \wedge & \text{discrete\_value(Patient, Timestamp, osmolality, Category)} \\ \wedge & \text{Category} \neq \text{high} \\ \longrightarrow & \neg\text{urine\_output\_reliable(Patient, Timestamp).} \end{array}$$

Eine Repräsentation der einzelnen Indikatoren als Menge von Fakten ist bereits für diese Teilmenge schwierig, da die letzte Regel eine Konjunktion und eine Prämisse „alle Kategorien bis auf ...“ enthält.
Es werden noch zahlreiche weitere Begriffe über Regelmengen abgeleitet, beispielsweise rapid_bp_fall/2, um einen rapide abfallenden Blutdruck anzuzeigen. prefer_dc_infu/2 drückt im Falle eines aktuell verabreichten Infusionsbolus eine Präferenz aus, diesen abzusetzen. tachycardia/2 gibt an, ob eine Herzfrequenz von größer 140 Schlägen pro Minute vorliegt. Die Argumente bezeichnen jeweils die Visite, auf die sich die Aussage bezieht[36].

3.6.6. Wahl der Therapiekandidaten

Die bisher beschriebenen Teile der Wissensbasis beschränken sich auf eine Ableitung von Begriffen, die einer abstrakteren Beschreibung des Patientenzustands dienen. Die Wahl der Therapien ist, wie schon im Abschnitt 3.4 erwähnt, der Teil des Behandlungsprotokolls mit den meisten Freiheitsgraden. Mit Hilfe der zuvor definierten Begriffe werden auf der hier beschriebenen Ebene Regeln formuliert, welche Therapien als Kandidaten in Frage kommen. Dabei dürfen sich Kandidaten zunächst noch widersprechen. Beispielsweise dürfen sich „Dosis eines Diuretikums erhöhen“ und „Dosis eines Diuretikums herabsetzen“ gleichzeitig in der näheren Wahl befinden. Die nachfolgenden Ebenen dienen dann der Auflösung von Therapie-Konflikten, der Einhaltung von Wartezeiten und der Ausarbeitung von Details zu Therapien, wie etwa die Berechnung von Dosierungen.

Vorkommende Therapien

Therapien entsprechen in der Wissensbasis Tupeln der Menge

<therapy_object> × <therapy_mode>,

wobei vom Mengenprodukt der Sorten nur eine Teilmenge verwendet wird. Die vorkommenden Therapien sind Tabelle 3.6 zu entnehmen.
Bei Infusionstherapien (Konstante: infusion) ist nicht nur die Zusammensetzung wichtig (vgl. Abschnitt 1.4). Schnell laufende Infusionen haben eine Auswirkung auf den Kreislauf. Die drei Modi zur Infusion stehen für unterschiedlich schnell laufende Infusionen. Bei einem Bolus (bolus) sollten laut Protokoll 15 ml/kg/h[37] in höchstens einer Stunde durchlaufen. Die Konstante maintenance bezeichnet Infusionen, die in „normaler“ Geschwindigkeit, bis zu 1.5 ml/kg/h laufen. Die Zusammensetzung richtet sich hierbei hauptsächlich nach der Natriumkonzentration im Blut, die zur Bewertung des Flüssigkeitshaushalts dient (vgl. Abschnitt 1.4.3). Die Konstante keep_open bezieht sich auf

[36] Also: predicate/2: <patient>, <time>.
[37] Die „kg“ beziehen sich auf das Körpergewicht.

<therapy_object>	<therapy_mode>
infusion	bolus maintenance keep_open
pressor	decrease increase urgent_increase
inotrope	decrease increase
diuretic	decrease increase

Tabelle 3.6.: Übersicht über die Therapien des Modells

Infusionen, die lediglich den Zugang über die Vene geöffnet halten sollen und sehr langsam laufen.
Neben Infusionstherapien bilden Veränderungen der Dosis eines Medikaments, das zu einer bestimmten Wirkungsgruppe gehört die zweite Art von Therapien. Die Medikamentengruppen entsprechen den in Tabelle 3.1 (S. 19) aufgeführten Wirkungsgruppen. Die Konstante inotrope bezeichnet positiv inotrope Medikamente, pressor entspricht der Gruppe der Vasokonstringenzien und diuretic bezeichnet Diuretika. In Notfallsituationen kann es vorkommen, dass Vasokonstringenzien schneller als gewöhnlich erhöht werden sollen. Aus diesem Grund existiert für diese Gruppe von Medikamenten, neben den Konstanten zur Dosissteigerung (increase) und -verminderung (decrease) eine weitere, für zügigere Dosiserhöhungen (urgent_increase).
Die auf dieser Ebene abgeleiteten Therapiekandidaten werden als Fakten des Prädikats

therapy_candidate/4: <patient>, <time>, <therapy_object>, <therapy_mode>.

repräsentiert.
Die Bedeutungen von Fakten, die Therapiekandidaten repräsentierten, lassen sich intuitiv anhand der Bezeichner ableiten. So drückt beispielsweise das Fakt

therapy_candidate(4999, 20, diuretic, increase).

aus, dass zum Zeitpunkt[38] 20 für Patient 4999 die Erhöhung der Dosis eines Diuretikums ein Therapiekandidat ist.

Regeln zur Ableitung von Kandidaten

Ähnlich, wie bei den komplexeren Abstraktionen und Bewertungen des Patientenzustands (vgl. Abschnitt 3.6.4) bieten sich auch bei der Ableitung von Therapiekandidaten kaum andere Repräsentationsformen an, als die Regeln unmittelbar in den MOBAL-Formalismus zu übertragen. Eine generellere Methode, die auf als Fakten repräsentiertes Sachbereichswissen zugreift, ist bei der vorliegenden Art von Wissen problematisch. Die Verschiedenartigkeit der Regeln, die unterschiedlichen beteiligten Prädikate und die uneinheitlichen Methoden der Herleitung verhindern eine derartige Abstraktion.

[38] Zeitpunkte werden in der Wissensbasis in der Einheit „Minuten seit Datenerfassung“ angegeben.

Die Ableitung der Therapiekandidaten ist recht umfangreich, weshalb im folgenden nur exemplarisch einige Aspekte aufgegriffen werden sollen.
Zu den wichtigsten Größen, an denen sich die Stabilität des Patientenzustands festmachen lässt, gehört der mittlere arterielle Blutdruck. Besonders wenn dieser zu niedrig ist, ist schnelles Handeln geboten. Die Anhebung des arteriellen Blutdrucks kann auf unterschiedliche Arten erreicht werden (vgl. S. 31). Das Behandlungsprotokoll sieht in diesem Fall Vasokonstringenzien und eine kreislaufstabilisierende Infusion vor. Die Wahl dieser Maßnahmen hängt davon ab, ob der mittlere arterielle Druck in die Kategorie threatening_low_bp fällt (vgl. S. 43). In diesem Fall wird zunächst unterschieden, ob es sich um einen Notfall handelt, der eine besonders zügige Gabe von Medikamenten indiziert, oder ob eine langsame Anhebung von Dosierungen ausreicht.
Ein Fakt des Prädikats threatening_hypotension/3 zeigt das Ergebnis dieser Bewertung an, wobei für die Unterscheidung zwischen Notfällen und weniger bedrohliche Situationen die Konstanten urgent und less_urgent zur Verfügung stehen.
Eine Einstufung als Notfall geschieht anhand der folgenden Regeln:

$$
\begin{array}{ll}
 & \mathsf{discrete\_value(P,\ T,\ mean\_arterial\_pressure,\ threatening\_low\_bp)} \\
\wedge & \mathsf{first\_iteration(P,\ T)} \\
\longrightarrow & \mathsf{threatening\_hypotension(P,\ T,\ urgent).}
\end{array}
$$

$$
\begin{array}{ll}
 & \mathsf{discrete\_value(P,\ T,\ mean\_arterial\_pressure,\ threatening\_low\_bp)} \\
\wedge & \mathsf{rapid\_bp\_fall(P,\ T)} \\
\wedge & \mathsf{significant\_change(P,\ T,\ central\_venous\_pressure,\ CVP\_Change)} \\
\wedge & \mathsf{CVP\_Change} \neq \mathsf{decrease} \\
\wedge & \neg\mathsf{high\_intracavitary\_pressure(P,\ T)} \\
\longrightarrow & \mathsf{threatening\_hypotension(P,\ T,\ urgent).}
\end{array}
$$

Notfälle sind laut Protokoll also gegeben, wenn ein extrem niedriger Blutdruck bereits zur ersten Visite (first_iteration/2) vorliegt.
Ist der niedrige arterielle Druck nach einem rapiden Abfall aufgetreten (rapid_bp_fall/2), der nicht mit einem gleichzeitigen Abfall des zentralvenösen Drucks einhergeht, und ist der intrathorakale Blutdruck[39] nicht zu hoch (negiertes high_intracavitary_pressure/2), so wird dies ebenfalls als Notfallsituation gewertet.
Die Menge aller Fälle bedrohlicher Hypotonie wird in den obigen Regeln anhand zusätzlich herangezogener Merkmale[40] partitioniert. Eine Partitionierung anhand weiterer Merkmale ist ein speziell bei der Therapiewahl häufig anzutreffendes Konzept. Es ist oft notwendig, alle möglichen Ergebnisse abfragen zu können. Im gezeigten Beispiel kommt die einfachste Methode zur Ableitung des „anderen Falles“ mit einer Anwendung des Operators „\+“ aus:

$$
\begin{array}{ll}
 & \mathsf{discrete\_value(P,\ T,\ mean\_arterial\_pressure,\ threatening\_low\_bp)} \\
\wedge & \mathsf{\backslash+(threatening\_hypotension(P,\ T,\ urgent))} \\
\longrightarrow & \mathsf{threatening\_hypotension(P,\ T,\ less\_urgent).}
\end{array}
$$

[39] Dieser Begriff bezeichnet entweder den zentralvenösen Blutdruck oder den pulmonalkappilären Verschlussdruck (vgl. S. 7) Letzterer hat Vorrang, ist aber selten verfügbar.
[40] Gemäß der vorgestellten Terminologie nach Puppe handelt es sich um „Merkmalsinterpretationen“.

Obwohl diese Repräsentation auch Nachteile mit sich bringt, ist die Regelmenge zur Ableitung von threatening_hypotension/3 nun redundanzfrei und dadurch wartbarer, übersichtlicher und weniger fehleranfällig, als wenn die beiden zuvor beschriebenen Regeln noch einmal in „invertierter Form“ in das Modell eingingen. Vor- und Nachteile autoepistemischer Operatoren werden in Abschnitt 3.8.2 noch ausführlicher diskutiert.

Nachdem die Art der bedrohlichen Hypotonie ermittelt wurde, werden die passenden Therapiekandidaten festgelegt. Ist der niedrige Blutdruck weniger bedrohlich, so wird

therapy_candidate(Patient, Timestamp, pressor, increase).

abgeleitet, Notfälle[41] führen zum Fakt

therapy_candidate(Patient, Timestamp, pressor, urgent_increase).

In Notfällen wird nun grundsätzlich die Gabe einer kreislaufstabilisierenden Infusion zum Therapiekandidaten. Im weniger dringenden Fall existieren hierfür zusätzliche Bedingungen. Die erste Bedingung ist, dass wirklich ein Medikament existiert, das im Rahmen der angedachten medikamentösen Therapie erhöht werden darf (drug_increaseable/3), weil es noch nicht in maximal zulässiger Dosierung gegeben wird. Ist dies der Fall, so leitet die Regel[42]

threatening_hypotension(Patient, Timestamp, less_urgent)
$\wedge$ pressor(Drug)
$\wedge$ drug_increaseable(Patient, Timestamp, Drug)
$\longrightarrow$ map_driven_infu(Patient, Timestamp).

ein Zwischenergebnis ab.

Die Struktur des Flussdiagramms ist in der Regel nicht baumartig, sondern es führen auf einer Seite oft mehrere Pfade zu einer Abfrage. Auch zu den Seiten selbst führen oft unterschiedliche Pfade. In der Wissensbasis werden für unterschiedlich erreichbare Stellen des Protokolls „Zwischenprädikate“ verwendet. Auch bei map_driven_infu/2 handelt es sich um solch ein Prädikat.

Gegenstand der nachfolgenden Ableitungsschritte ist die Untersuchung weiterer Merkmale, die eventuell gegen eine schnell laufende Infusion sprechen. Eine solche Infusion erhöht das Füllungsvolumen des Kreislaufs, das aber bereits recht hoch sein kann. Dies wird anhand des intrathorakalen Blutdrucks abgeschätzt. Ist es tatsächlich zu hoch, so wird bei einer unzureichenden arteriellen Durchblutung statt dessen die Gabe eines inotropen Medikaments in Betracht gezogen. Neben der Bewertung der arteriellen Durchblutung gehen weitere komplexe Zwischenergebnisse der Abstraktionsebene in die Ableitungen ein.

Je nach den Ergebnissen der Auswertung wird schließlich entweder einer der beiden Therapiekandidaten (infusion, bolus) und (inotrope, increase), oder kein weiterer Therapiekandidat abgeleitet.

Auf die einzelnen Regeln soll aufgrund der Komplexität nicht weiter eingegangen werden.

[41] Eigentlich hängt dies noch von der zusätzlichen Bedingung ab, dass der systolische Blutdruck zu niedrig ist. Ein über mehrere Minuten ohne Medikamentenerhöhung stabiler systolischer Druck führt zum Abbruch der Notfallbehandlung, die in gewisser Weise von normalen Visiten abweicht. Dies wird in Abschnitt 3.6.11 genauer ausgeführt.

[42] Dabei testet pressor/1, ob ein Medikament zu den im Rahmen der Therapie (pressor, increase) erhöhten Medikamente gehört.

Bemerkungen zur Grundlage der Therapiewahl

In Abschnitt 3.4 wurden die Regeln zur Therapiewahl im Protokoll als heuristische Assoziationen bezeichnet. Anhand der Beispiele in diesem Abschnitt lässt sich nun gut nachvollziehen, dass die Herleitung nicht etwa über repräsentierte kausale Zusammenhänge geschieht, sondern dass von Verknüpfungen von Merkmalsinterpretationen[43] unmittelbar auf Therapiekandidaten geschlossen wird. Die in dieser Diplomarbeit ausgeführten medizinischen Zusammenhänge sind größtenteils nicht im Protokoll expliziert, sondern wurden in Gesprächen mit einem Experten und aus Fachliteratur zusammengetragen.

Versucht man die Plausibilität der in diesem Abschnitt vorgestellten Schlüsse ohne solche Zusatzinformationen nachzuvollziehen, so fällt die mangelnde Verankerung der Schlüsse durch expliziertes Sachbereichswissen auf. Auch über den in diesem Abschnitt beschriebenen Modellabschnitt hinaus ist selbst die Wirkung der Medikamente nicht Teil des Modells. Lediglich eine Zuordnung der Medikamente zu Begriffen wie „Vasokonstriktor“ kann als gegeben betrachtet werden. Dementsprechend kann bereits vorweggenommen werden, dass die über heuristische Assoziationen gewonnen Therapiekandidaten auch an späterer Stelle nicht mehr anhand von sicherem Wissen gegengeprüft werden, wie dies etwa in der Architektur aus [MIB+00] vorgesehen ist (vgl. auch Abschnitt 1.5.1).

3.6.7. Konfliktauflösung für sich ausschließende Therapien

Nach der Beschreibung der Herleitung von Therapiekandidaten im letzten Abschnitt geht es in diesem Abschnitt darum, aus der Menge eventuell unvereinbarer Kandidaten eine konsistente Teilmenge auszuwählen.

Zur Repräsentation unvereinbarer Therapiekandidaten dient das Prädikat

conflicting_therapies/6: <patient>, <time>, <therapy_object>, <therapy_mode>, <therapy_object>, <therapy_mode>.

Das dritte und vierte Argument geben den ersten Kandidaten an, die letzten beiden den zweiten. Die folgenden Regeln definieren die Unvereinbarkeit von Kandidaten im Sinne des Behandlungsprotokolls:

$$
\begin{array}{ll}
 & \text{therapy\_candidate(Patient, Time, Obj, Mode1)} \\
\wedge & \text{therapy\_candidate(Patient, Time, Obj, Mode2)} \\
\wedge & \text{Mode1} \neq \text{Mode2} \\
\longrightarrow & \text{conflicting\_therapies(Patient, Time, Obj, Mode1, Obj, Mode2).}
\end{array}
$$

Zu jeder Visite ist also nur einer der Infusionsmodi bolus, maintenance und keep_open zulässig. Ebenso dürfen zu jeder Medikamentengruppe nicht gleichzeitig increase und decrease oder increase und urgent_increase akzeptiert werden.

Die folgende Regel bezieht sich ausschließlich auf Dosisverminderungen.

[43] Verknüpfungen von Merkmalsinterpretationen können selbst wiederum als Merkmalsinterpretationen aufgefasst werden.

$$
\begin{array}{ll}
 & \neg\mathsf{rapid\_rx\_dec(Patient,\ Time)} \\
\wedge & \mathsf{therapy\_candidate(Patient,\ Time,\ Obj1,\ decrease)} \\
\wedge & \mathsf{therapy\_candidate(Patient,\ Time,\ Obj2,\ decrease)} \\
\wedge & \mathsf{Obj1} \neq \mathsf{Obj2} \\
\longrightarrow & \mathsf{conflicting\_therapies(Patient,\ Time,\ Obj1,\ decrease,\ Obj2,\ decrease).}
\end{array}
$$

Es darf also pro Visite nur für maximal eine Wirkungsgruppe die Verminderung einer Medikamentendosis abgeleitet werden. Das erste Literal der Prämisse schließt die Anwendbarkeit dieser Regel in Fällen aus, in denen aufgrund eines zu hohen arteriellen Blutdrucks Medikamente schneller als gewöhnlich abgesetzt werden dürfen. In solchen Fällen dürfen Medikamente auch „parallel“ abgesetzt werden.

Grundlage der nächsten Regel ist, dass das Protokoll in Notfällen[44] auf zwei feste Therapien beschränkt ist, (pressor, urgent_increase) und (infusion, bolus). Der erste Therapiekandidat wird ausschließlich in Notfällen abgeleitet, so dass folgende Regel formuliert werden kann:

$$
\begin{array}{ll}
 & \mathsf{therapy\_candidate(Pat,\ Time,\ pressor,\ urgent\_increase)} \\
\wedge & \mathsf{therapy\_candidate(Pat,\ Time,\ Obj,\ Mode)} \\
\wedge & \mathsf{Mode} \neq \mathsf{urgent\_increase} \\
\wedge & \mathsf{Mode} \neq \mathsf{bolus} \\
\longrightarrow & \mathsf{conflicting\_therapies(Pat,\ Time,\ pressor,\ urgent\_increase,\ Obj,\ Mode).}
\end{array}
$$

Schließlich nutzt die letzte Regel die Symmetrie der Konfliktrelation aus, um den Kontrollregeln letztlich alle im Konflikt liegenden geordneten Therapiepaare in Faktenform zur Verfügung zu stellen:

$$
\begin{array}{ll}
 & \mathsf{conflicting\_therapies(Patient,\ Time,\ Obj1,\ Mode1,\ Obj2,\ Mode2)} \\
\longrightarrow & \mathsf{conflicting\_therapies(Patient,\ Time,\ Obj2,\ Mode2,\ Obj1,\ Mode1).}
\end{array}
$$

Bevor auf die Regeln zur Auflösung von Konflikten eingegangen wird, soll zunächst die verwendete Ordnung über der im letzten Abschnitt bereits vorgestellten Menge möglicher Therapien vorgestellt werden. Diese ist anhand von Fakten des Prädikats

therapy_hierarchy/3: <therapy_object>, <therapy_mode>, <therapy_id>.

definiert. Die ersten beiden Argumente jedes Fakts bezeichnen gemeinsam einen Therapiekandidaten. Diesem wird jeweils eine eineindeutige Priorität (<therapy_id>) zugewiesen.

Die Auflösung von Konflikten geschieht mit Hilfe der so definierten Ordnung. Liegen zwei Therapiekandidaten im Konflikt, so wird der Kandidat niedrigerer Priorität zurückgewiesen. Tabelle 3.7 zeigt die verwendete Hierarchie.

Das Flussdiagramm enthält keine tabellarische oder ähnliche Darstellung dieser Hierarchie. Einige Therapie-Anweisungen lassen bereits den Vorrang anderer Therapien erkennen, und bestimmte widersprüchliche Therapiekandidaten werden nie zusammen abgeleitet. Einige Konflikte werden im Protokoll auch über das Prinzip zeitlichen Vorrangs

[44] Mit „Notfällen“ sind immer schwere hypotone Krisen gemeint, da das Protokoll nur für diese Art von Notfällen eine gesonderte Behandlung vorsieht.

<therapy_object>	<therapy_mode>	<therapy_id>
pressor	urgent_increase	9
pressor	increase	8
infusion	bolus	7
pressor	decrease	6
inotrope	increase	5
inotrope	decrease	4
infusion	maintenance	3
diuretic	increase	2
diuretic	decrease	1
infusion	keep_open	0

Tabelle 3.7.: Vorranghierarchie der Therapien

aufgelöst. Die Änderung von Medikamentendosierungen wird verhindert, falls kurz zuvor (insbesondere in der gleichen Visite) bereits eine andere, unverträgliche Änderung durchgeführt wurde. Um die Hierarchie explizieren zu können, war daher insbesondere zu erfassen, welche Therapien beim Durchlaufen des Flussdiagramms vor welchen anderen durchgeführt werden.

Die gewählte Form der Modellierung ist im Gegensatz dazu deklarativ. Bei zwei unvereinbaren Therapiekandidaten wird der Kandidat geringerer Priorität zurückgewiesen. Bezüglich des produzierten Verhaltens entsprechen sich Flussdiagramm und Modell, jedoch ist die Hierarchie des Modells übersichtlicher und leichter zu warten.

Die folgenden zwei Regeln operationalisieren die oben beschriebene Strategie zur Auflösung von Konflikten:

```
     conflicting_therapies(Patient, Time, Obj1, Mode1, Obj2, Mode2)
∧    therapy_hierarchy(Obj1, Mode1, Priority1)
∧    therapy_hierarchy(Obj2, Mode2, Priority2)
∧    Priority1 < Priority2
∧    therapy(Patient, Time, Obj2, Mode2)
⟶    reject_therapy(Patient, Time, Obj1, Mode1).

     therapy_candidate(Patient, Time, Obj, Mode)
∧    \+(reject_therapy(Patient, Time, Obj, Mode))
⟶    therapy(Patient, Time, Obj, Mode).
```

Akzeptierte Therapien werden als Fakten des Prädikats

therapy/4: <patient>, <time>, <therapy_object>, <therapy_mode>.

repräsentiert. In der Verfeinerungsphase werden solche Therapien durch die Zuweisung konkreter Medikamente bzw. Infusionen und Dosierungen bzw. Volumen zu Interventionsvorschlägen konkretisiert.

3.6.8. Therapieverfeinerungen

Nachdem die auszuführenden Therapien feststehen, wird aus Lesbarkeitsgründen zunächst eine Übersetzung aus dem Ausgangsformat vorgenommen, bei dem den meisten Therapien eigene Prädikate zugeordnet werden, beispielsweise

therapy(Patient, Timestamp, inotrope, decrease)
$\longrightarrow$ decrease_inotropes(Patient, Timestamp).

Die darauf aufbauenden Inferenzen konkretisieren die zunächst nur grob beschriebenen Therapien. Dabei werden keinerlei Therapiekandidaten mehr abgeleitet. Weiterhin bestehen keine Abhängigkeiten zwischen parallel ausgeführten Therapien, sondern die Verfeinerungen der Therapien sind rein lokaler Natur.
Bei Therapien, die sich auf Wirkungsgruppen von Medikamenten beziehen, ist der erste verfeinernde Schritt die Wahl eines konkreten Medikaments. Bei Infusionstherapien ist eingangs nur der Modus bekannt. Der erste Schritt ist daher die Wahl der Infusion selbst.

Wahl von Medikamenten zu Wirkungsgruppen

In den vorigen Abschnitten[45] wurde bereits die Auswahl anhand von Hierarchien als eine typische Methode des Behandlungsprotokolls vorgestellt. Dabei wird jeweils aus einer mit Prioritäten annotierten Menge der zulässige Eintrag höchster Priorität ausgewählt. Auch für die Wahl eines Medikaments zu einer abstrakten Therapie sind Hierarchien zentral. Dabei variieren die Auswahlstrategien von „hierarchisch unter Einbeziehung weiterer Merkmale“ über „rein hierarchisch“ bis zu „Auswahl des einzigen zur Gruppe gehörenden Medikaments“.
Zunächst seien die Zielprädikate dieser Ebene vorgestellt. Die zugehörigen Fakten enthalten neben den Visitendaten das ausgewählte Medikament. Die Bezeichner der Prädikate weisen auf die durchzuführende Aktion hin. Die Ausführung geschieht auf einer späteren Ebene. Bei

drug_step_down/3: <patient>, <time>, <drug>.

soll die Dosis des angegebenen Medikaments verringert werden.
Heraufgesetzt werden soll die Dosis des angegebenen Medikaments bei

drug_step_up/3: <patient>, <time>, <drug>.

Fakten des Prädikats

drug_stop_overdose/3: <patient>, <time>, <drug>.

werden im Fall einer vorliegenden und zu beseitigenden Überdosis inferiert. Überdosierte Medikamente werden nicht grundsätzlich herabgesetzt, sondern dies muss ausdrücklich abgeleitet werden. Allerdings ist zu sagen, dass es bei der konsequenten Anwendung des Protokolls nicht zu Überdosierungen kommen kann.
Soll die Dosis eines Vasokonstriktors heraufgesetzt werden, so wird bei der Wahl des Medikaments nicht zwischen der normalen und der Notfallvariante unterschieden. Dementsprechend fällt die Unterscheidung bei der Ableitung des Aktionsprädikats zunächst weg.

[45] Beispielsweise geschieht auf Seite 49 ff. die Bewertung der Veränderung der arteriellen Durchblutung hierarchisch.

```
        therapy(Patient, Timestamp, pressor, increase)
  ⟶    increase_pressors(Patient, Timestamp).

        therapy(Patient, Timestamp, pressor, urgent_increase)
  ⟶    increase_pressors(Patient, Timestamp).
```

Bei der Wahl wird auf die folgende Hierarchie zurückgegriffen:

```
pressor_hierarchy( dopamin,  1 ).
pressor_hierarchy( arteren,  2 ).
pressor_hierarchy( supra,    3 ).
```

Die Eigenschaft, vasokonstringent zu wirken, wurde in den vorigen Abschnitten als dosisabhängig definiert. Die Zuordnungen finden sich in Tabelle 3.1 auf Seite 19. Beim Erhöhen dieser Medikamente wird jedoch keine derartige Unterscheidung vorgenommen[46]. Alle im Protokoll vorkommenden Medikamente, die in einer beliebigen Dosis vasokonstringent wirken, sind aufgeführt. Es wird einfach das erste Medikament (kleinste ID) gewählt, das noch unterhalb der maximalen Dosierung gegeben wird. Dazu werden zunächst mit Hilfe eines Zwischenprädikats die Medikamente „aufgesammelt", die den Bedingungen genügen, also in der Hierarchie vertreten und noch zu erhöhen sind:

```
        increase_pressors(Patient, Time)
  ∧    pressor_hierarchy(Drug, Drug_ID)
  ∧    drug_increaseable(Patient, Time, Drug)
  ⟶    inc_pressor_candidate(Patient, Time, Drug_ID).
```

Mit Hilfe des autoepistemischen Operators min_of wird aus dieser Menge das Medikament mit der geringsten ID ermittelt, dessen Dosis dann per drug_step_up/3 erhöht wird:

```
        increase_pressors(Patient, Time)
  ∧    min_of(ID, inc_pressor_candidate(Patient, Time, ID), Drug_ID)
  ∧    pressor_hierarchy(Drug, Drug_ID)
  ⟶    drug_step_up(Patient, Time, Drug).
```

Sollen Vasokonstringenzien abgesetzt werden, so geschieht dies wiederum mit Hilfe der beschriebenen Hierarchie, wobei die Medikamente in der umgekehrten Reihenfolge herab- wie heraufgesetzt werden:

```
        decrease_pressors(Patient, Time)
  ∧    max_of(ID, dec_pressor_candidate(Patient, Time, ID, M), Drug_ID)
  ∧    pressor_hierarchy(Drug, Drug_ID)
  ⟶    drug_step_down(Patient, Time, Drug).
```

[46] Andernfalls würde sich beispielsweise Adrenalin (Suprarenin), das laut Tabelle erst ab einer mittleren Dosis vasokonstringent wirkt, gar nicht langsam bis in diesen Dosisbereich erhöhen lassen.

Die Auswahl der Kandidaten, die schließlich durch Fakten des Prädikats

dec_pressor_candidate/4: <patient>, <time>, <pressor_id>, <inotr_mode>.

repräsentiert werden, ist allerdings komplexer, als beim Heraufsetzen. Die Medikamente wirken teilweise in bestimmten Dosisbereichen auch positiv inotrop, beschleunigen also die Herzfrequenz. Wenn die Herzfrequenz nun bereits zu hoch ist (Tachykardie), dann setzt das Protokoll bevorzugt einen Vasokonstriktor herab, der in der aktuellen Dosis gleichzeitig positiv inotrop wirkt. Ist die Herzfrequenz bekanntermaßen nicht zu hoch, so werden bevorzugt nicht zusätzlich noch positiv inotrop wirkende Medikamente abgesetzt. Gibt es in einem der Fälle kein derartiges Medikament, oder ist die Herzfrequenz unbekannt, so werden alle Vasokonstringenzien zu Kandidaten. Das Prädikat tachycardia/2 zeigt durch positive Literale eine bekannte Tachykardie an, durch negative hingegen, dass bekanntermaßen keine Tachykardie vorliegt. Ist die Herzfrequenz unbekannt, so wird nichts abgeleitet[47].
Die folgenden Regeln leiten bei bekannter Herzfrequenz die bevorzugten Kandidaten her:

```
   decrease_pressors(Patient, Time)
∧  tachycardia(Patient, Time)
∧  drug_effect_category(Patient, Time, pressor, Drug)
∧  drug_effect_category(Patient, Time, inotrope, Drug)
∧  pressor_hierarchy(Drug, Drug_ID)
⟶  dec_pressors_candidate(Patient, Time, Drug_ID, inotrope_check).
```

```
   decrease_pressors(Patient, Time)
∧  ¬tachycardia(Patient, Time)
∧  drug_effect_category(Patient, Time, pressor, Drug)
∧  \+(drug_effect_category(Patient, Time, inotrope, Drug))
∧  pressor_hierarchy(Drug, Drug_ID)
⟶  dec_pressors_candidate(Patient, Time, Drug_ID, inotrope_check).
```

Die folgende Regel deckt sowohl den Fall ab, dass kein Medikament mit den bevorzugten Eigenschaften gefunden wurde, als auch den einer unbekannten Herzfrequenz, da dann die beiden vorigen Regeln nicht anwendbar sind:

```
   decrease_pressors(Pat, Time)
∧  \+(dec_pressors_candidate(Pat, Time, ID, inotrope_check))
∧  drug_effect_category(Pat, Time, pressor, Drug)
∧  pressor_hierarchy(Drug, Drug_ID)
⟶  dec_pressors_candidate(Pat, Time, Drug_ID, no_inotrope_check).
```

Einfacher gestaltet sich die Änderung der positiv inotropen Medikamente. Es wird streng hierarchisch das zu ändernde Medikament ermittelt[48]. Beim Erhöhen ist die einzige zusätzliche Bedingung, dass noch nicht die maximal zulässige Dosis verabreicht wird. Wird

[47] Dies entspricht dem Beispiel von S. 21 ff.

[48] Dopamin wirkt in mittlerer Dosis positiv inotrop (vgl. Tabelle 3.1), wird aufgrund der an dieser Stelle sehr deutlichen Protokollanweisungen jedoch nicht zusammen mit den anderen inotropen Medikamenten erhöht. Beim Herabsetzen wird Dopamin allerdings wie die übrigen inotropen Medikamente behandelt.

beim Testen dieser Bedingung ein überdosiertes Medikament bemerkt, so wird dies mittels drug_stop_overdose/3 auf die größte zulässige Dosis herabgesetzt. Die Dosis des gewählten Medikaments wird über drug_step_up/3 heraufgesetzt. Analog wird beim Herabsetzen das gemäß Hierarchie erste Medikament mit einer Dosis größer Null gewählt. Zusätzliche Bedingung ist hierbei jedoch eine positiv inotrope Wirkung der aktuellen Dosis.

Die Hierarchie für die Wirkungsgruppe der positiv inotropen Medikamente wird durch die nachfolgenden Fakten definiert, wobei beim Heraufsetzen Medikamente mit niedrigerem, beim Herabsetzen Medikamente mit höherem annotierten Wert bevorzugt werden.

```
inotrope_hierarchy( dobutrex,   1 ).
inotrope_hierarchy( dopamin,    2 ).
inotrope_hierarchy( supra,      3 ).
```

Das Herabsetzen inotroper Medikamente geschieht wie folgt:

```
      decrease_inotropes(Patient, Time)
  ∧   drug_effect_category(Patient, Time, inotrope, Drug)
  ∧   inotrope_hierarchy(Drug, Drug_ID)
 ⟶   dec_inotropes_candidate(Patient, Time, Drug_ID).

      decrease_inotropes(Patient, Time)
  ∧   max_of(ID, dec_inotropes_candidate(Patient, Time, ID), Drug_ID)
  ∧   inotrope_hierarchy(Drug, Drug_ID)
 ⟶   drug_step_down(Patient, Time, Drug).
```

Bei Diuretika ist die Wahl eines konkreten Medikaments besonders leicht, da das Protokoll nur das Medikament Lasix verwendet. Es wird allerdings unterschieden, ob das Diuretikum kontinuierlich oder als Bolus verabreicht wird[49]. Im Behandlungsprotokoll ist Lasix das einzige Medikament, für das auch eine Bolusgabe vorgesehen ist. Die Verabreichungsform ist besonders im Falle einer fortgeführten Behandlung wichtig. Wurde vor wenigen Stunden noch Lasix gegeben, so wird der Modus beibehalten, da die Dosierung sich dann nach den vorangegangenen Dosierungen richtet. Wurde Lasix hingegen schon mehrere Stunden lang nicht mehr verwendet, so spezifiziert das Protokoll die Verabreichungsform auch nicht weiter, so dass mangels Bedeutung dieses Details in der Umsetzung einfach eine kontinuierliche Gabe vorgeschlagen wird.

Zum Herabsetzen von Diuretika ist zu sagen, dass dies nur für die kontinuierliche Verabreichung von Lasix Bedeutung hat, da Bolusgaben nur kurze Zeit andauern. Das Herauf- und Herabsetzen der Dosis diuretischer Medikamente geschieht – trotz einiger Besonderheiten – ebenfalls über drug_step_up/3 und drug_step_down/3.

Da bei medikamentenbasierten Therapien weder gewährleistet ist, dass überhaupt ein passendes Medikament gefunden wird, noch, dass nur eine Therapie akzeptiert wird, existieren für diese Aspekte des Modells keine Constraints.

[49]Ein Bolus ist durch eine einmalige schnelle Gabe eines Wirkstoffs gekennzeichnet. Im gegebenen Fall würde also eine bestimmte Menge „auf einmal" gespritzt. Im Gegensatz dazu können Wirkstoffe auch über einen bestimmten Zeitraum kontinuierlich zugeführt werden.

Wahl einer konkreten Infusion

Ähnlich wie bei den auf Medikamenten basierenden Therapien werden auf dieser Ebene auch die Infusions-Bestandteile zu den Infusionstherapien gewählt. Das Prädikat, mit dem die gewählten Infusionen repräsentiert werden, ist

chosen_infusion/4: <patient>, <time>, <infusion>, <infusion_mode>.

Neben den Visitendaten und dem Modus zur Gabe der Infusion besitzt jedes Fakt eine Konstante der Sorte <infusion>, die die Zusammensetzung der Infusion bestimmt. Die Sorte enthält die Konstanten ringer (Ringer-Laktat Lösung), d5w (Glukose 5%) und tutofusin (Tutofusin OPX). Die Bezeichner sind dabei, in Vorbereitung auf nachfolgende Experimente, an hierzulande verwendete Infusionen angelehnt.
Für die beiden Infusionsmodi keep_open und bolus kann einfach eine Ringer-Laktat Lösung gewählt werden. Beim Infusionsmodus maintenance richtet sich hingegen die Zusammensetzung nach dem Flüssigkeitshaushalt des Körpers (vgl. Abschnitt 1.4.3). Dieser wird hauptsächlich anhand der Natriumkonzentration des Blutes (effective_na[50]) bewertet. Die passende Infusion wird daher anhand der Kategorie dieses Wertes gewählt:

$$\begin{array}{rl} & \text{infu\_hydration\_based(Patient, Time)} \\ \wedge & \text{discrete\_value(Patient, Time, effective\_na, Category)} \\ \wedge & \text{hydration\_to\_infusion(Category, Infusion)} \\ \longrightarrow & \text{chosen\_infusion(Patient, Time, Infusion, maintenance).} \end{array}$$

Die Zuordnung von Kategorien zu Infusionen ist den folgenden Fakten zu entnehmen:

```
hydration_to_infusion( under_hydration,    d5w ).
hydration_to_infusion( normal_hydration,   tutofusin ).
hydration_to_infusion( over_hydration,     ringer ).
hydration_to_infusion( unknown,            tutofusin ).
```

Eine Rahmenbedingung ist die Existenz und Eindeutigkeit einer Konkretisierung zu jeder akzeptierten Infusionstherapie. Da gleichzeitig pro Visite maximal eine Infusionstherapie akzeptiert wird, können die folgenden Constraints formuliert werden:

$$\begin{array}{rl} & \text{therapy(Patient, Timestamp, infusion, Mode)} \\ \leadsto & \text{chosen\_infusion(Patient, Timestamp, I, Mode).} \end{array}$$

$$\begin{array}{rl} & \text{chosen\_infusion(Patient, Timestamp, Infu1, Mode1)} \\ \wedge & \text{chosen\_infusion(Patient, Timestamp, Infu2, Mode2)} \\ \leadsto & (\text{Infu1} = \text{Infu2}) \wedge (\text{Mode1} = \text{Mode2}). \end{array}$$

3.6.9. Wartezeiten

Im Behandlungsprotokoll werden Änderungen an der Dosis eines Medikaments teilweise aufgrund von bestehenden Wartezeiten nicht durchgeführt. Neben der Ablehnung von

[50] Die Variable effective_na setzt sich genau genommen aus der Natriumkonzentration und dem Blutzuckerspiegel zusammen, wobei letzterer nur bei zu hohen Werten Berücksichtigung findet.

Therapiekandidaten im Konfliktfall stellt dies den zweiten Grund dar, aus dem eine zunächst gewählte Therapie eventuell nicht ausgeführt wird. Zunächst seien zwei Arten von Wartezeiten im Protokoll unterschieden.
Die erste Art von Wartezeiten legt die Zeitspanne fest, die mindestens zwischen zwei Erhöhungen oder zwei Verminderungen der Dosis des gleichen Medikaments liegen muss. Diese Art von Wartezeit ist hier nicht gemeint. Sie wird in Abschnitt 3.6.10 zusammen mit der Dosisermittlung behandelt.
Die zweite Art von Wartezeiten wird aufgrund oszillierender Dosisverläufe festgesetzt. Diese sind als Dosiserhöhung kurz nach einer Verminderung definiert. Ein oszillierender Dosisverlauf kann als erfolgloser Versuch gewertet werden, eine medikamentöse Unterstützung zurückzuziehen. Der nächste Versuch das Medikament herabzusetzen wird dann durch eine Wartezeit verzögert. Die Voraussetzungen zum Setzen einer Wartezeit sind der folgenden Regel zu entnehmen:

```
     visit(Patient, Time)
∧    drug_effect_category(Patient, Time, pressor, P)
∧    drug_can_have_waiting_time(Drug)
∧    new_drug_oscill(Patient, Time, Drug, NumberOf, LastOscillsInc)
⟶    set_waiting_time(Patient, Time, Drug).
```

Es muss also zur Visite ein Vasokonstriktor verabreicht werden, für das betrachtete Medikament müssen Wartezeiten vorgesehen sein (drug_can_have_waiting_time/1), und seit der letzten Visite muss ein oszillierender Dosisverlauf des Medikaments vorliegen (new_drug_oscill/5). Die Gabe eines Vasokonstriktors ist offenbar deshalb Bedingung, weil die Wartezeiten nur für den Fall massiver kreislaufstützender Maßnahmen gedacht sind.
Die Länge gesetzter Wartezeiten richtet sich nach der Anzahl der oszillierenden Dosisverläufe des Medikaments innerhalb eines Zeitfensters fester Größe. Bestehende Wartezeiten werden durch Fakten des Prädikats

waiting_time/4: <patient>, <time>, <drug>, <time>.

repräsentiert. Nach den Angaben zur Visite folgt das Medikament, auf das sich die Wartezeit bezieht, und der Zeitpunkt, bis zu dem gewartet werden soll. Zur Herleitung wird das Prädikat oscills_to_waittime/2 verwendet, das eine Abbildung der Häufigkeit des „Oszillier-Ereignisses“ auf Wartezeiten in Minuten beschreibt[51]:

```
     set_waiting_time(Patient, Time, Drug)
∧    new_drug_oscill(Patient, Time, Drug, NumberOf, LastOscillsInc)
∧    oscills_to_waittime(NumberOf, Minutes_To_Wait)
∧    add(LastOscillsInc, Minutes_To_Wait, Wait_Until)
∧    Wait_Until > Time
⟶    waiting_time(Patient, Time, Drug, Wait_Until).
```

Die Variable NumberOf bezeichnet in der obigen Regel die Zahl der Schwankungen. Genauer ist dies die Anzahl der Dosisverminderungen innerhalb der letzten 24 Stunden, für

[51]Dabei führt das MOBAL built-in add/3: <arg1>, <arg2>, <arg3>. die Berechnung $Arg_3 := Arg_1 + Arg_2$ durch.

die die darauf folgende Dosisveränderung eine Erhöhung war, die maximal 3 Stunden später statt fand. Die Variable LastOscillsInc bezeichnet den Zeitpunkt der Dosiserhöhung, die zum letzten beobachteten Oszillieren der Dosis des betrachteten Medikaments gehört. Dieser Zeitpunkt wird auch für die Berechnung des Zeitpunktes benötigt, bis zu dem die Wartezeit in Kraft bleiben soll.
Da die zum Prädikat new_drug_oscill/5 gehörenden Ereignisse sich schlecht unmittelbar in MOBAL auswerten lassen, erfolgt die Auswertung mit Hilfe des externen Prädikats

ask_oscillating(+Patient, +Time, +Drug, +ITS, +OTS, -NumberOf, -LastOscillsInc).

(vgl. S. 38), wobei die 3 und 24 Stunden als Parameter in Fakten abgelegt sind:

```
    visit(Pat, Time)
∧   oscill_data(inner_timewindow_size, ITS)
∧   oscill_data(outer_timewindow_size, OTS)
∧   drug_attrib(Drug)
∧   attribute_mapping(DB, Drug)
∧   extern(ask_oscillating(Pat, Time, DB, ITS, OTS, NumberOf, LastInc))
⟶   oscills_before_visit(Pat, Time, Drug, NumberOf, LastInc).

    oscill_data(inner_timewindow_size, 180).
    oscill_data(outer_timewindow_size, 1440).
```

Wenn das Oszillieren nach der Vorgängervisite auftritt, so wird zum oscills_before_visit/5-Fakt das passende new_drug_oscill/5-Fakt abgeleitet. Andernfalls wird die Wartezeit nicht neu berechnet, da dies bereits in der letzten Visite geschehen ist. Darüber hinaus können Wartezeiten auch abgebrochen werden, was bedingt, dass nachfolgende Visiten diese nicht lokal neu berechnen. Abbrüche von Wartezeiten werden durch ein wie folgt abgeleitetes Fakt des Prädikats cancel_waiting_time/3 angezeigt:

```
    drug_step_down(Patient, Time, Drug)
∧   waiting_time(Patient, Time, Drug, Wait_Until)
∧   rapid_rx_dec(Patient, Time)
⟶   cancel_waiting_time(Patient, Time, Drug).
```

Bedingung ist hier also ein dringender Bedarf, Medikamente abzusetzen. Dieser wird an einem normalen oder zu hohen arteriellen Blutdruck festgemacht, was erneut die überwiegende Ausrichtung des Mechanismus „Wartezeiten“ auf kreislaufstabilisierende Medikamente verdeutlicht. Die einzige zusätzliche Bedingung ist, dass versucht wird, das betroffene Medikament trotz aktueller Wartezeit abzusetzen.
Aktuell vorliegende, nicht abgebrochene Wartezeiten werden als Fakten des Prädikats

still_waiting_time/3: <patient>, <time>, <drug>.

repräsentiert. Wird innerhalb einer Visite für ein Medikament sowohl eine Wartezeit, als auch ein Abbruch der Wartezeit abgeleitet, so gilt die Wartezeit insgesamt als abgesetzt. Dies ergibt sich aus der Auswertungsreihenfolge im Protokoll. Ob eine aktuelle Wartezeit besteht, wertet die nachfolgende Regel aus.

$$
\begin{array}{rl}
 & \text{waiting\_time(Patient, Time, Drug, Wait\_Until)} \\
\wedge & \text{\textbackslash+(cancel\_waiting\_time(Patient, Time, Drug))} \\
\longrightarrow & \text{still\_waiting\_time(Patient, Time, Drug).}
\end{array}
$$

Schließlich müssen noch die aus vorigen Visiten fortbestehenden Wartezeiten auf die jeweils aktuelle Visite übertragen werden:

$$
\begin{array}{rl}
 & \text{visit(Patient, Now)} \\
\wedge & \text{next\_visit(Patient, LastVisit, Now)} \\
\wedge & \text{still\_waiting\_time(Patient, LastVisit, Drug)} \\
\wedge & \text{max\_of(T, waiting\_time(Patient, LastVisit, Drug, T), Wait\_Until)} \\
\wedge & \text{Wait\_Until} > \text{Now} \\
\longrightarrow & \text{waiting\_time(Patient, Now, Drug, Wait\_Until).}
\end{array}
$$

Das Prädikat next_visit/3 verkettet dabei aufeinander folgende Visiten. Wartezeiten werden nur übertragen, falls sie noch gültig sind. Existieren für ein Medikament mehrere Einträge für Zeitpunkte, bis zu denen gewartet werden soll, so wird nur der letzte Zeitpunkt (max_of) betrachtet. Dieser muss nach dem Visitenzeitpunkt liegen.

3.6.10. Ableitung konkreter Interventionsvorschläge

Die letzten Schritte bis zum Interventionsvorschlag umfassen bei den medikamentenbasierten Therapien neben Berechnungen zur Dosis auch Tests, ob Wartezeiten verletzt werden. Bei Infusionstherapien wird das anzuhängende Volumen berechnet.

Bestimmung der Dosis

Die Ausgangsdaten zur Berechnung der Medikamentendosierung sind Fakten der Prädikate drug_step_up/3 und drug_step_down/3 (vgl. Abschnitt 3.6.8). Die Berechnungen hängen vom jeweiligen Medikament ab. Ein Überschreiten der maximal zulässigen Dosis ist in keinem der Fälle vorgesehen. Heraufgesetzt wird immer mindestens auf die minimale Dosis gemäß drug_dose_bounds/3 (vgl. Seite 44). Entsteht beim Herabsetzen eine Dosis unterhalb der Minimaldosis, so wird das Medikament ganz abgesetzt. Im Behandlungsprotokoll werden bezüglich der Modalitäten zur Dosisänderung drei Arten von Medikamenten unterschieden.

Bei der ersten Gruppe wird die Dosis jeweils linear erhöht und vermindert. Innerhalb der zulässigen Dosisgrenzen ergibt sich die neue Dosis also durch Addition oder Subtraktion einer vom Medikament abhängigen Konstanten.

Bei der zweiten Gruppe können die Medikamente beim Heraufsetzen nur bestimmte Dosierungen annehmen. Beispielsweise darf Dobutrex laut Protokoll nur auf 5, 10, 15 und 20 μg/kg/min heraufgesetzt werden. Herabgesetzt werden diese Medikamente hingegen linear. Hierbei können auch beim Heraufsetzen unzulässige Dosierungen auftreten.

Schließlich gibt es noch Medikamente, deren Dosis beim Heraufsetzen jeweils verdoppelt wird. Herabgesetzt werden auch diese linear, abgesehen von Lasix als Bolus, bei dem ein Absetzen aber ohnehin weder nötig noch möglich ist[52].

Die Wissensbasis enthält eine über Fakten realisierte Zuordnung der Medikamente zu Änderungsmodi. Für das Herauf- und Herabsetzen existieren jeweils eigene Einträge, um

[52] Ein Bolus wird annähernd „auf einmal“ gegeben.

Änderungen zu erleichtern. Die folgenden Fakten geben die Zuordnung wieder:

```
drug_change_mode( arteren,        increase,  linear ).
drug_change_mode( supra,          increase,  linear ).
drug_change_mode( phenylephrine,  increase,  linear ).
drug_change_mode( dopamin,        increase,  levelwise ).
drug_change_mode( dobutrex,       increase,  levelwise ).
drug_change_mode( lasix_drip,     increase,  double_dose ).
drug_change_mode( lasix_bolus,    increase,  double_dose ).
drug_change_mode( lasix_bolus,    decrease,  none ).
drug_change_mode( arteren,        decrease,  linear ).
drug_change_mode( ...             decrease,  linear ).
...                               ...        ...
```

Die Dosierungen der übrigen Medikamente werden linear herabgesetzt. Im Modus none wird die weitere Verarbeitung einfach abgebrochen. Die Bedeutung der Konstanten ist den Bezeichnern zu entnehmen.

Die folgenden Constraints stellen sicher, dass zu jedem Medikament und jeder Änderungsrichtung genau ein Modus der Dosisänderung existiert:

```
   drug_attrib(Drug)
⇝  drug_change_mode(Drug, increase, Mode).

   drug_attrib(Drug)
⇝  drug_change_mode(Drug, decrease, Mode).

   drug_change_mode(Drug, Direction, Mode1)
∧  drug_change_mode(Drug, Direction, Mode2)
⇝  Mode1 = Mode2.
```

Die Regeln zur Berechnung der Dosisänderung sind einfach, aufgrund der vielen Fälle aber unübersichtlich. Im folgenden wird exemplarisch der Fall einer linearen Dosiserhöhung vorgestellt. Zunächst wird mit Hilfe der obigen Zuordnung ein Prädikat definiert, das die durchzuführende Aktion und den Änderungsmodus zusammenfasst:

```
    drug_step_up(Patient, Timestamp, Drug)
∧   drug_change_mode(Drug, increase, linear)
∧   drug_increaseable(Patient, Timestamp, Drug)
⟶   increase_dose_linear(Patient, Timestamp, Drug).
```

Die vorgesehene neue Dosis wird nun in Abhängigkeit von der diskreten Kategorie[53] der aktuellen Dosis festgesetzt. Ist die Dosis nicht bekannt, so wird eine Dosis von Null angenommen[54]. Die folgende Regel dient dem Hochsetzen auf die minimale Dosis, falls

[53] Diskrete Dosierungen werden anhand des Prädikats discrete_drug_dose/4 repräsentiert, vgl. Abschnitt 3.6.4.

[54] Eine unbekannte Dosis bedeutet bei der verwendeten Interpretation der Daten, dass bis zum fraglichen Zeitpunkt kein einziger Eintrag vorliegt. Dies konnte beim in dieser Arbeit verwendeten Datensatz so interpretiert werden, dass das entsprechende Medikament noch nicht gegeben wurde.

die aktuelle Dosis weder im Normalbereich noch darüber liegt:

increase_dose_linear(Patient, Timestamp, Drug)
$\wedge$ discrete_drug_dose(Patient, Timestamp, Drug, Category)
$\wedge$ Category $\neq$ normal_dose
$\wedge$ Category $\neq$ over_dose
$\wedge$ drug_dose_bounds(Drug, LowerBound, UpperBound)
$\longrightarrow$ check_drug_dose(Patient, Timestamp, Drug, LowerBound, increase).

Das Prädikat check_drug_dose/5 repräsentiert Kandidaten für Interventionsvorschläge, die nur noch auf die Verletzung von Wartezeiten hin überprüft werden müssen.
Im normalen Dosisbereich wird nun – in Einklang mit der Definition linearer Dosiserhöhungen – eine vom Medikament abhängige Konstante zur aktuellen Dosis addiert. Um eine gleichartige Repräsentation der verschiedenen Fälle zu erreichen, hängt die Wahl der Dosisschrittweite noch von der Dringlichkeit der Dosisänderung ab, die zu jeder Visite anhand des Prädikats

dose_change_speed/4: <patient>, <time>, <drug_change_dir>, <change_speed>.

abgelesen werden kann. Die Sorte <drug_change_dir> enthält dabei die Konstanten increase und decrease, die Sorte <change_speed> die Konstanten normal und fast. Der Therapiemodus urgent_increase geht mit einer schnelleren Dosiserhöhung einher, ansonsten wird mit normaler Geschwindigkeit erhöht. Ebenso führt ein rapid_rx_dec/2 während einer Visite zu einer schnelleren Verminderung, andernfalls wird eine normale Änderungsgeschwindigkeit abgeleitet.

Die Schrittweiten für die lineare Erhöhung der Dosis von Medikamenten werden als Fakten des Prädikats

drugstep_value/4: <drug>, <change_speed>, <step_value>.

repräsentiert, wobei über Constraints sichergestellt ist, dass zu jedem Tupel aus linear erhöhtem Medikament und Änderungsgeschwindigkeit eine eindeutige Schrittweite existiert.
Um nach der Addition der Konstanten keine Überdosis zu erhalten, wird schließlich das Minimum[55] aus maximal zulässiger und errechneter Dosis verwendet.
Die entsprechende Regel aus dem Modell lautet

increase_dose_linear(Patient, Timestamp, Drug)
$\wedge$ discrete_drug_dose(Patient, Timestamp, Drug, normal_dose)
$\wedge$ dose_change_speed(Patient, Timestamp, increase, Speed)
$\wedge$ drugstep_value(Drug, Speed, Diff)
$\wedge$ actual_value(Patient, Timestamp, Drug, Dose, Value_From)
$\wedge$ add(Dose, Diff, Sum)
$\wedge$ drug_dose_bounds(Drug, LowerBound, UpperBound)
$\wedge$ extern(minimum(Sum, UpperBound, NewDose))
$\longrightarrow$ check_dose_change(Patient, Timestamp, Drug, NewDose, increase).

[55] Die Berechnung von Minimum und Maximum wurde über externe Prädikate „minimum/3“ und „maximum/3“ realisiert. Das dritte Argument enthält dabei jeweils das Ergebnis.

Im Rahmen der meisten Dosisänderungen wird als Zwischenschritt ein Fakt des Prädikats check_dose_change/5 abgeleitet, das die Informationen eines vollständigen Interventionsvorschlag beinhaltet und die nun näher zu betrachtenden Tests auf Wartezeiten anstößt.

Überprüfen von Wartezeiten

Neben der in Abschnitt 3.6.9 beschriebenen Art von Wartezeiten gibt es im Protokoll noch generelle Mindestwartezeiten zwischen Dosisänderungen. Beide Arten beziehen sich auf konkrete Medikamente, nicht auf Wirkungsgruppen. Die Anwendung hängt zum Teil von der aktuellen Dosis, stärker aber noch von der Änderungsrichtung ab. Wird für eine beabsichtigte Änderung die Verletzung einer Wartezeit festgestellt, so wird ein entsprechendes Fakt des Prädikats

cancel_propose/4: <patient>, <time>, <drug>, <drug_change_dir>.

abgeleitet. Dies drückt aus, dass die vorgesehene Dosisänderung für das Medikament (<drug>) in die angegebene Richtung (<drug_change_dir>) nicht durchgeführt werden kann. Im folgenden werden die Regeln vorgestellt, die dazu führen, dass Vorschläge unterbunden werden.
Die erste Regel testet die Mindestwartezeiten zwischen zwei Änderungen des gleichen Medikaments in die gleiche Richtung. Sie gilt sowohl beim Erhöhen, als auch beim Vermindern der Dosis. Darüber hinaus werden die unterschiedlichen Mindestwartezeiten zu den Änderungsgeschwindigkeiten normal und fast berücksichtigt.

$$
\begin{array}{ll}
 & \mathsf{check\_dose\_change(Patient,\ Timestamp,\ Drug,\ Dose,\ Direction)} \\
\wedge & \mathsf{discrete\_drug\_dose(Patient,\ Timestamp,\ Drug,\ Category)} \\
\wedge & \mathsf{Category} \neq \mathsf{none} \\
\wedge & \mathsf{Category} \neq \mathsf{under\_dose} \\
\wedge & \mathsf{dose\_change\_speed(Patient,\ Timestamp,\ Direction,\ Speed)} \\
\wedge & \mathsf{mintime\_between\_changes(Direction,\ Speed,\ MinDiff)} \\
\wedge & \mathsf{last\_dose\_change(Patient,\ Timestamp,\ Drug,\ Direction,\ When)} \\
\wedge & \mathsf{sub(Timestamp,\ When,\ TimeDiff)} \\
\wedge & \mathsf{TimeDiff} < \mathsf{MinDiff} \\
\longrightarrow & \mathsf{cancel\_propose(Patient,\ Timestamp,\ Drug,\ Direction).}
\end{array}
$$

Die zugehörigen Mindestwartezeiten des Protokolls in Minuten sind den folgenden Fakten zu entnehmen:

```
mintime_between_changes( increase,  fast,    3 ).
mintime_between_changes( increase,  normal,  5 ).
mintime_between_changes( decrease,  fast,    5 ).
mintime_between_changes( decrease,  normal,  60 ).
```

Die Regel schließt die Dosiskategorien none und under_dose aus, da für aktuelle Dosierungen unterhalb der Mindestdosis keine Wartezeiten vorgesehen sind[56]. Die Dosis wird in diesem Fall entweder auf die Mindestdosis erhöht, oder das Medikament wird abgesetzt.

[56] Wenn die Kategorie unknown wie none interpretiert wird, wie es im Modell der Fall ist, dann gilt dies sinngemäß auch für diese Kategorie. Allerdings liegen in diesem Fall gar keine Werte zur Dosis vor und die Mindestwartezeit kann nicht verletzt sein.

Als zusätzliche Beschränkung darf bei einer Dosisverminderung mit normaler Geschwindigkeit für das Medikament keine der in Abschnitt 3.6.9 vorgestellten Wartezeiten vorliegen:

```
     check_dose_change(Patient, Timestamp, Drug, NewDose, decrease)
∧    NewDose > 0.0
∧    dose_change_speed(Patient, Timestamp, decrease, normal)
∧    still_waiting_time(Patient, Timestamp, Drug)
⟶    cancel_propose(Patient, Timestamp, Drug, decrease).
```

Ein Vorschlag an den Anwender wird genau dann abgeleitet, wenn einer in Betracht gezogenen Intervention keine Wartezeit im Wege steht:

```
     check_dose_change(Patient, Timestamp, Drug, Dose, Direction)
∧    \+(cancel_propose(Patient, Timestamp, Drug, Direction))
⟶    propose_drug(Patient, Timestamp, Drug, Dose, Direction).
```

Das Prädikat

propose_drug/5: <patient>, <time>, <drug>, <dose>, <drug_change_dir>.

dient der Repräsentation medikamentöser Interventionsvorschläge. Zu Zwecken der Übersichtlichkeit enthalten die zugehörigen Fakten dabei neben der neuen Dosis auch die Richtung der vorgeschlagenen Änderung.
Zur Interpretation durch menschliche Anwender sind noch die Einheiten der Dosierungen wichtig. Arterenol, Dobutrex, Dopamin, Suprarenin und Phenylephrine sind in der Einheit [μg/kg/min] angegeben. Dabei bezieht sich das „kg“ auf das Körpergewicht. Die Einheit der Dosierungen kontinuierlicher Lasix-Gaben (Konstante: lasix_drip) ist [mg/kg/min]. Die Konstante lasix_bolus bezeichnet eine einmalige Gabe des Medikaments Lasix. Die Einheit ist hierbei [mg].

Berechnung des Infusionsvolumens

Bei der Ableitung der Interventionsvorschläge zu Infusiontherapien genügt es, für die bereits hergeleiteten Fakten des Prädikats chosen_infusion/4 das anzuhängende Infusionsvolumen zu ergänzen. Die Zusammensetzung und die Durchlaufgeschwindigkeit der Infusion stehen bereits fest. Wartezeiten, wie bei den Medikamenten, bestehen nicht. Gründe, die gegen eine bestimmte Infusionstherapie sprechen, werden bereits bei der Therapiewahl oder Konfliktauflösung berücksichtigt.
Die resultierenden Interventionsvorschläge werden mit Hilfe des Prädikats

propose_infusion/5: <patient>, <time>, <infusion>, <infusion_ml>,
<infusion_mode>.

ausgedrückt. Die Wahl des Volumens gestaltet sich wie folgt:
Soll lediglich der Zugang auf gehalten werden (keep_open), so kann ein geringes konstantes Volumen von 10 ml/h angegeben werden.
Bei Infusionen zur Kreislaufstabilisierung (bolus) richtet sich das anzuhängende Volumen nach dem Körpergewicht. Es ergibt sich als Produkt des Gewichts mit dem Faktor 15 ml/kg/h.

Bei der letzten Art von Infusionen (maintenance) hängt das Volumen noch von weiteren Patientenvariablen ab. Im Falle einer Dialyse oder eines viel zu hohen Creatinin-Wertes wird als Faktor nur 0.5 ml/kg/h zugrundegelegt, andernfalls 1.5 ml/kg/h. Das vorgeschlagene Volumen ergibt sich daraus wieder durch Multiplikation mit dem Körpergewicht.
Während bei den Infusionsmodi keep_open und bolus das letzte Argument der abgeleiteten Fakten einfach das anzuhängende Volumen für eine Stunde (Einheit: [ml/h]) angibt, gibt der Wert bei maintenance die angestrebte *Gesamt*flüssigkeitszufuhr einer Stunde an. Das Infusionsvolumen ergibt sich somit als Differenz zur übrigen aufgenommenen Flüssigkeit.

3.6.11. Der Visitenzyklus

Die in den letzten Abschnitten beschriebenen Teile des Modells sind an die Inferenzstruktur der heuristischen Klassifikation angelehnt. Sie beziehen sich auf einzelne Visiten, für die die benötigten Daten erhoben, mit Referenzdaten verglichen und abstrahiert werden. Schließlich werden in einem mehrstufigen Prozess anhand der Daten abstrakte Therapien gewählt, die dann verfeinert und schließlich vorgeschlagen werden.
Das Behandlungsprotokoll sieht eine zyklische Wiederholung von Visiten vor. Wie bereits im Abschnitt 3.6.9 kurz erwähnt, sind aufeinander folgende Visiten miteinander verkettet und die Inferenzen zur späteren Visite hängen teilweise von der vorigen Visite ab. Beispielsweise werden noch gültige Wartezeiten für Medikamente aufgrund oszillierender Dosierungen auf diese Weise fortgeführt. Die Ermittlung von Referenzwerten (Abschnitt 3.6.3) geschieht allerdings nicht anhand der in vorigen Visiten erhobenen Daten, da die Protokollvorgaben durch direkte Zugriffe auf die Datenschnittstelle genauer erfüllt werden.
Jede Visite wird durch ein Fakt des Prädikats

visit/2: <patient>, <time>.

angestoßen. Handelt es sich um die erste Visite, so wird ein first_iteration/2-Fakt abgeleitet. Das gleiche gilt auch für vereinzelt angestoßene Visiten[57], zu denen kein Vorgänger bekannt ist. Die Vorgänger-Relation für Visiten wird über das Prädikat

next_visit/3: <patient>, <time>, <time>.

definiert. Der zweite Zeitpunkt ist dabei jeweils unmittelbarer Nachfolger des ersten. Vorgänger und Nachfolger zu Visiten müssen eindeutig sein. Die folgenden Constraints überwachen die Anforderungen an die Relation:

$$\begin{array}{ll} & \text{next\_visit(Patient, T1, T2)} \\ \leadsto & \text{T1} < \text{T2.} \end{array}$$

$$\begin{array}{ll} & \text{next\_visit(Patient, Time, Next1)} \\ \wedge & \text{next\_visit(Patient, Time, Next2)} \\ \leadsto & \text{Next1 = Next2.} \end{array}$$

[57] Dies lässt sich in MOBAL einfach durch Einfügen eines visit/2-Faktes erreichen.

$$\begin{array}{ll} & \text{next\_visit(Patient, Before1, Time)} \\ \wedge & \text{next\_visit(Patient, Before2, Time)} \\ \rightsquigarrow & \text{Before1 = Before2.} \end{array}$$

Während jeder Visite wird der Zeitpunkt der jeweils nächsten Visite festgesetzt. Dies geschieht durch die Ableitung eines entsprechenden next_visit/3-Faktes. Die Zeitspanne, die dabei zwischen zwei Visiten liegt, hängt vom Zustand des Patienten ab. Das Behandlungsprotokoll unterscheidet zunächst einmal drei solcher Zeitspannen, 30 Minuten, 2 Stunden und 4 Stunden. Die Zuordnung zu Patientenzuständen ist in der vorliegenden Fassung jedoch nicht formalisiert. Da es sich offenbar um Obergrenzen handelt, und da im Protokoll auftretende zeitliche Bedingungen ohnehin ausdrücklich getestet werden[58], wird in der Umsetzung im normalen Ablauf standardmäßig eine Zeitspanne von 30 Minuten gewählt.

Neben den im Protokoll ausdrücklich benannten Zeitspannen gibt es noch eine weitere, die sich aus dem Verhalten in Notfallsituationen ergibt. Tritt eine lebensbedrohliche Hypotonie auf, so wird nicht der übliche Visitenzyklus fortgesetzt, sondern es wird ausschließlich und kontinuierlich versucht, den Patienten zu stabilisieren. Hierbei wird alle drei Minuten ein Medikament heraufgesetzt, so lange, bis die Abbruchbedingung erfüllt ist. Die Behandlung von Notfällen ist in der Umsetzung in die bereits beschriebene Struktur der heuristischen Klassifikation eingebunden. Es werden Therapiekandidaten generiert, die gemäß der üblichen Konfliktbehandlung akzeptiert werden. Es wird ein zu erhöhendes Medikament gewählt, wobei die einzige in diesem Fall geprüfte Wartezeit 3 Minuten beträgt. Auch die Abwicklung über den Visitenzyklus findet bei Notfällen Anwendung, allerdings dann mit einer Zeitspanne von ebenfalls nur 3 Minuten zwischen den einzelnen Visiten.

Notfälle werden anhand des Prädikates emergency_state/2 angezeigt. Neu auftretende Notfälle hängen mit lebensbedrohlicher Hypotonie zusammen. Die in Abschnitt 3.6.6 bereits beschriebene Ableitung des Therapiekandidaten (pressor, urgent_increase)[59] im Falle einer solchen hypotonen Krise leitet als Zwischenschritt ein „Notfallfakt“ ab:

$$\begin{array}{ll} & \text{threatening\_hypotension(Patient, Timestamp, urgent)} \\ \longrightarrow & \text{emergency\_state(Patient, Timestamp).} \end{array}$$

$$\begin{array}{ll} & \text{emergency\_state(Patient, Timestamp)} \\ \longrightarrow & \text{therapy\_candidate(Patient, Timestamp, infusion, bolus).} \end{array}$$

$$\begin{array}{ll} & \text{emergency\_state(Patient, Timestamp)} \\ \longrightarrow & \text{therapy\_candidate(Patient, Timestamp, pressor, urgent\_increase).} \end{array}$$

Die Therapiekandidaten sind somit eine generelle Konsequenz von Notfallsituationen. Letztere werden, außer über die erste der obigen Regeln, auch noch anhand der nachfolgenden Regel abgeleitet.

[58] Zu den Tests im Protokoll gehört beispielsweise die Mindestwartezeit von nur 5 Minuten zwischen zwei Dosiserhöhungen des gleichen Medikaments.

[59] „Erhöhe zügig die Dosis eines Vasokonstriktors!“

$$
\begin{array}{ll}
 & \mathsf{visit(Patient,\ Now)} \\
\wedge & \mathsf{next\_visit(Patient,\ Last\_Visit,\ Now)} \\
\wedge & \mathsf{emergency\_state(Patient,\ Last\_Visit)} \\
\wedge & \mathsf{\backslash+(stop\_emergency\_protocol(Patient,\ Last\_Visit))} \\
\longrightarrow & \mathsf{emergency\_state(Patient,\ Now).}
\end{array}
$$

Der Notfallstatus überträgt sich also von einer Visite zur nächsten, es sei denn, das Abbruchkriterium ist erfüllt. Abgebrochen wird erst, wenn der systolische Druck seit mindestens 6 Minuten nicht mehr unter dem Normalintervall liegt, ohne dass Medikamente heraufgesetzt wurden. Das Prädikat stable_systolic_bp_since/3 gibt den Zeitpunkt an, zu dem der Wert das letzte Mal unterhalb des Normalintervalls lag. Liegt der aktuellste Wert unterhalb, so wird kein Fakt abgeleitet. Das Abbruchkriterium lässt sich wie folgt formulieren:

$$
\begin{array}{ll}
 & \mathsf{emergency\_state(Pat,\ Now)} \\
\wedge & \mathsf{stable\_systolic\_bp\_since(Pat,\ Now,\ BPSince)} \\
\wedge & \mathsf{max\_of(T,\ last\_dose\_change(Pat,\ Now,\ Drug,\ increase,\ T),\ DSince)} \\
\wedge & \mathsf{extern(maximum(BPSince,\ DSince,\ Stable\_Since))} \\
\wedge & \mathsf{sub(Now,\ Stable\_Since,\ Timespan)} \\
\wedge & \mathsf{Timespan} \geq 6 \\
\longrightarrow & \mathsf{stop\_emergency\_protocol(Pat,\ Now).}
\end{array}
$$

Bei stabilem systolischen Blutdruck wird in Notfällen bereits die passende Therapie (pressor, urgent_increase) zusammen mit allen anderen Therapien, die die Dosis von Medikamenten ändern zurückgewiesen. Dies ist notwendig, damit das Abbruchkriterium überhaupt erfüllt werden kann, bevor alle Medikamente in voller Dosierung gegeben werden[60].

Nachdem die Integration von Notfällen in das Visitenkonzept dargestellt wurde, bleibt noch die Ableitung der jeweils nächsten Visite und die benutzerseitige Steuerung der Visiten zu zeigen.
In Abhängigkeit von der Visitenart wird die passende Zeitspanne zur nächsten Visite gewählt. Falls kein Notfall vorliegt, so geschieht das beispielsweise anhand der Regel

$$
\begin{array}{ll}
 & \mathsf{visit(Patient,\ Now)} \\
\wedge & \neg\mathsf{emergency\_state(Patient,\ Now)} \\
\wedge & \mathsf{timespan\_between\_iterations(normal\_protocol,\ Interval)} \\
\wedge & \mathsf{add(Now,\ Interval,\ Time\_Next\_Visit)} \\
\longrightarrow & \mathsf{next\_visit(Patient,\ Now,\ Time\_Next\_Visit).}
\end{array}
$$

wobei timespan_between_iterations/2 eine Abbildung der Visitenmodi auf Zeitspannen realisiert. Ob zum next_visit/3-Fakt eine Visite angestoßen wird hängt davon ab, ob der Zeitpunkt der nächsten Visite in einem benutzerseitig angegebenen Intervall liegt, für das Visiten durchgeführt werden sollen.

Das Prädikat für die Angabe solcher Intervalle ist

visit_from_to/3: <patient>, <time>, <time>.

[60] Wenn alle 3 Minuten ein Medikament hochgesetzt wird, dann ist durchgängig die Bedingung verletzt, 6 Minuten lang keine Medikamente heraufgesetzt zu haben.

Die Zeitpunkte beschreiben den Zeitpunkt der ersten Visite und den Zeitpunkt, ab dem keine weiteren Visiten mehr durchgeführt werden sollen. Um den Visitenzyklus für ein Intervall anzustoßen und anhand des Modells Interventionsvorschläge abzuleiten genügt es, ein solches Fakt in MOBAL einzutragen. Die folgenden Regeln steuern den Ablauf:

$$
\begin{array}{ll}
 & \mathsf{visit\_from\_to(Patient,\ Start,\ End)} \\
\longrightarrow & \mathsf{visit(Patient,\ Start).}
\end{array}
$$

$$
\begin{array}{ll}
 & \mathsf{next\_visit(Patient,\ Now,\ Time\_Next\_Visit)} \\
\wedge & \mathsf{visit\_from\_to(Patient,\ Start,\ End)} \\
\wedge & \mathsf{Time\_Next\_Visit} \geq \mathsf{Start} \\
\wedge & \mathsf{Time\_Next\_Visit} \leq \mathsf{End} \\
\longrightarrow & \mathsf{visit(Patient,\ Time\_Next\_Visit).}
\end{array}
$$

Es können auf diese Weise auch für mehrere Intervalle Visitenzyklen durchlaufen werden, allerdings sollten diese möglichst disjunkt sein, damit die Eindeutigkeitsbedingungen zu next_visit/3 erfüllt bleiben.

3.7. Die Wissensbasis an einem Beispiel

In diesem Abschnitt soll kurz illustriert werden, wie sich die Arbeit mit der Wissensbasis in MOBAL gestaltet. Über die in Abschnitt 3.6.11 beschriebenen Mechanismen lassen sich Visiten einzeln oder für ein Intervall im Sinne des Visitenzyklus verkettet anstoßen. In Abhängigkeit von den Patientendaten werden daraufhin abstraktere Therapien und konkrete Interventionen vorgeschlagen. Um nun eine Ableitung nachvollziehen zu können, stehen in MOBAL geeignete Hilfsmittel zur Verfügung.

Zunächst einmal können alle abgeleiteten Fakten in MOBAL-Fenstern betrachtet werden. Diese verfügen über eine Fokus-Funktion, die etwa eine Selektion von Fakten eines bestimmten Patienten, Zeitpunktes und Prädikats ermöglicht.

Abbildung 3.4 zeigt einige Originalfenster einer beispielhaften Visite. Auf die Herkunft der Daten wird in Abschnitt 4.1 näher eingegangen. Die aufgeführten Fakten stellen nur einen Auszug für die betrachtete Visite dar, bieten aber bereits einen Einblick in verschiedene Abstraktionsebenen[61].

Das erste Fenster zeigt die unterste Abstraktionsebene, auf der die aktuellsten Daten zur Visite anhand des Prädikats **actual_value/5** dargestellt werden. In diskretisierter Form finden sich diese Daten dann in den Fenstern zu den Prädikaten **discrete_value/4** und **discrete_drug_dose/4**. Anhand von **unstability_criteria/3** und **drug_effect_category/4** lassen sich weitere Erkenntnisse über den aktuellen Patientenzustand gewinnen. Die Veränderung der Messwerte zeigt das Prädikat **significant_change/4** an. Schließlich lassen sich noch die vom Protokoll abgeleiteten Therapiekandidaten ablesen.

Die Daten können von Experten dazu verwendet werden, den vorliegenden Patientenzustand sowie die Angemessenheit der vom Protokoll vorgeschlagenen Therapie zu bewerten.

[61] Zu actual_value/5 siehe Abschnitt 3.6.2, zu discrete_value/4, discrete_drug_dose/4 und drug_effect_category/4 Abschnitt 3.6.4. In Abschnitt 3.6.5 findet sich eine Beschreibung zu unstability_criteria/3, und in Abschnitt 3.6.6 zu therapy_candidate/4.

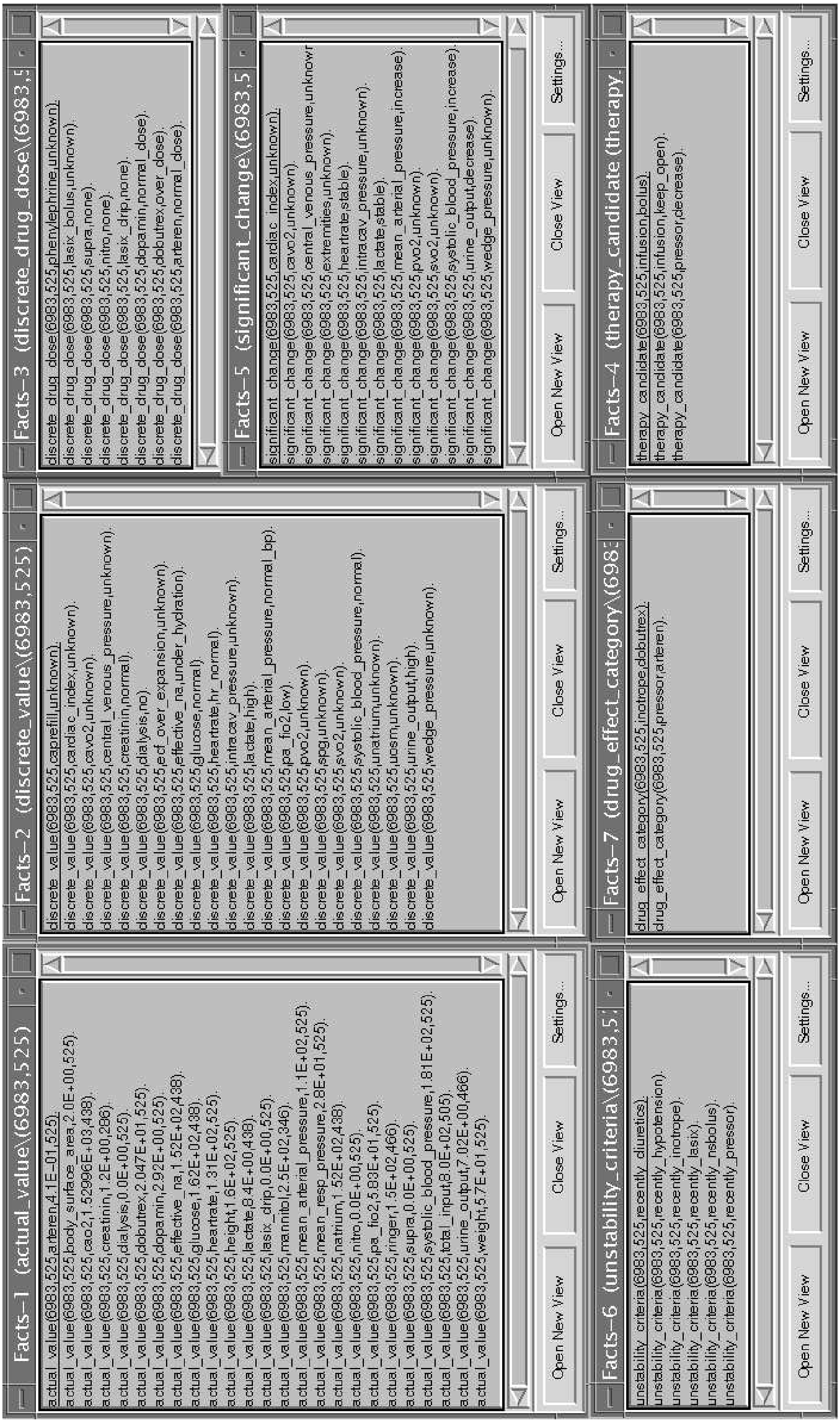

Abbildung 3.4.: Fakten zur Visite beim Patienten 6983 zum Zeitpunkt 525

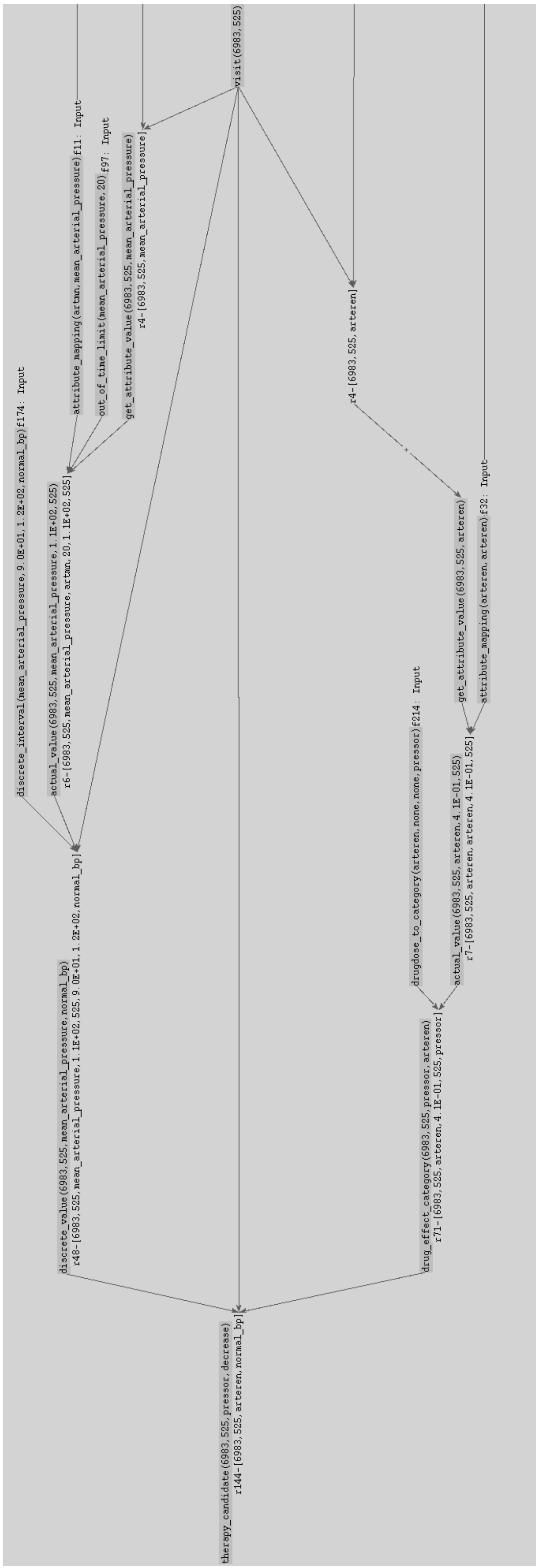

Abbildung 3.5.: Auszug aus dem Ableitungsgraphen zu einem Therapiekandidaten

Ist nun etwa die Ableitung eines Therapiekandidaten aus medizinischer Sicht unverständlich, so kann MOBAL den entsprechenden Ableitungsgraphen zu diesem Kandidaten anzeigen. Abbildung 3.5 zeigt einen unvollständigen Graphen zu einem Kandidaten aus Abbildung 3.4. Zu jedem einzelnen Schritt kann die entsprechende Regel der Wissensbasis angezeigt werden. Auf diese Weise lassen sich Zusammenhänge schnell überblicken und Fehler leichter finden.

3.8. Abschließende Bemerkungen zum Modell

3.8.1. Die Behandlung von Zeit

Die Architektur der in diesem Kapitel beschriebenen Wissensbasis ist an die Inferenzstruktur der heuristischen Klassifikation angelehnt. Diese Sicht beschreibt vorwiegend einzelne Visiten. Die Behandlung von Zeit wird erst auf spezifischeren Ebenen sichtbar. In diesem Kapitel soll die Behandlung von Zeit im Protokoll noch einmal zusammenfassend beschrieben werden. Die in [SM96] beschriebene Arbeit beschäftigt sich umfassend mit zeitlichen Abstraktionen im klinischen Umfeld. Der Ansatz ist im Sinne der Newellschen Wissensebene (siehe [New82]) Domänen-unabhängig und daher recht universell einsetzbar. Die meisten Aspekte von Zeitverarbeitung im Protokoll lassen sich in diesen Kontext einordnen.

Die zeitlichen Abläufe des Protokolls werden durch den Visitenzyklus bestimmt. Visiten beziehen sich in der Repräsentation auf einzelne Zeitpunkte, die relevanten Daten entstammen jedoch aus der unmittelbaren, teils auch aus der länger zurückliegenden Vergangenheit.

Bei der Erhebung der jeweils aktuellen Werte lassen sich grundsätzlich zwei Arten von Variablen unterscheiden. Bei der ersten Art können Werte so lange als exakt angenommen werden, bis ein neuer Wert vorliegt. Als Beispiel seien festgelegte Dosierungen kontinuierlich gegebener Medikamente genannt. Veralterung von Werten gibt es in diesem Zusammenhang nicht. Die Gültigkeit von Werten der zweiten Sorte ist hingegen zeitlich begrenzt. Hierzu zählen besonders erhobene Messwerte wie der Herzindex oder Blutzuckerspiegel. In Abhängigkeit von der Variablen-Art sind die betrachteten zeitlichen Intervalle demnach entweder beidseitig oder nur nach oben beschränkt [62]. Die maximale Zeitspanne, für die ein erhobener Wert verwendet werden darf, gehört zum Domänen-Wissen und wird in [SM96] als *temporal persistance* bezeichnet. Die Ableitung der jeweils aktuellsten Werte zu jeder Visite stellt eine erste Abstraktion dar, die als Zeitintervall-zu-Zeitpunkt Projektion aufgefasst werden kann. In diese Kategorie fällt weiterhin die Suche nach bestimmten Ereignissen wie Hypotonie, über längere Zeitintervalle. Dies dient im Protokoll beispielsweise der Stabilitätsbewertung von Patienten.

Viele der Variablen werden nicht unmittelbar in ihrer numerischen Ausgangsform, sondern diskretisiert verwendet. Dies entspricht der *vertical temporal inference* aus [SM96], die dort recht allgemein als Klassifikationsaufgabe beschrieben wird. Die Ergebnisse sind dabei vom Abstraktionstyp „Zustand“ (*state*).

Die Bewertung von Parameter-Veränderungen im Protokoll entspricht einem anderen Abstraktionstyp (*gradient*). Die Änderung wird über die absolute Differenz des aktuellsten Wertes mit einem älteren Referenzwert bewertet. Für die Bewertung von Veränderungen als signifikant oder nicht signifikant wird wiederum Domänen-Wissen benötigt. Die

[62] Genauer haben die Intervalle für einen Zeitpunkt t_1, einen Visitenzeitpunkt t_2 und Startzeitpunkt 0 entweder die Form $(t_1, t_2]$ oder $[0, t_2]$.

Differenz, ab der von einer signifikanten Veränderung gesprochen werden kann, wird in [SM96] als Beispiel für den Begriff des *temporal dynamic knowledge* verwendet.
Beim Test, ob Dosisverläufe oszillieren, sieht das Protokoll eine Suche nach zeitlichen Mustern vor (*temporal pattern matching* in [SM96]). Werden solche Muster entdeckt, so kann dies zu Wartezeiten führen, in denen Medikamente nicht abgesetzt werden dürfen. Definiert ist ein oszillierender Dosisverlauf als Dosiserhöhung kurz nach einer Dosisverminderung. Für welchen Zeitraum eine Wartezeit angesetzt wird, hängt davon ab, wie viele solcher Ereignisse bereits in einem größeren Zeitfenster stattgefunden haben. Wartezeiten können bei gesteigertem Bedarf auch vorzeitig abgesetzt werden. Sie stellen eine Art von strategischem Element dar.
Zum Begriff des „Auswertungskontextes“ aus [SM96] ist anzumerken, dass im Protokoll Abstraktionen immer globalen Charakter haben, also etwa die *vertical temporal inference* oder gar die *temporal persistance* nicht an spezifische Kontexte gebunden sind.

3.8.2. Verwendung autoepistemischer Operatoren

In diesem Abschnitt sollen einige Schwierigkeiten erläutert werden, die sich bei der Umsetzen des Behandlungsprotokolls in den von MOBAL verwendeten Formalismus ergaben. Der Schwerpunkt liegt auf einer Diskussion der Vor- und Nachteile autoepistemischer Operatoren, die im Modell mehrfach vorkommen, obwohl sie zu Problemen im Zusammenspiel mit einigen MOBAL-Werkzeugen führen.

Hierarchien

Ein wiederholt im Behandlungsprotokoll auftretendes Muster ist die Auswahl anhand einer Hierarchie. Im ersten Schritt werden dabei anhand bestimmter Kriterien und Beschränkungen Kandidaten generiert. Im zweiten Schritt wird aus der Kandidatenmenge dann das hierarchiehöchste Element gewählt. Beispiele für dieses Vorgehen finden sich u.a. bei der Abstraktion von Merkmalen, etwa auf Seite 49 ff., wo die Änderung der arteriellen Durchblutung bewertet wird, oder auch bei der Medikamentenwahl, die in Abschnitt 3.6.8 näher beschrieben wird.
Um unabhängig von der konkreten Ausprägung der Hierarchie die Wahl eines hierarchiehöchsten Elements umzusetzen, sind autoepistemische Operatoren ein unverzichtbares Hilfsmittel. Die Operatoren `min_of` bzw. `max_of` können aus den Kandidatenmengen unmittelbar den gesuchten Kandidaten ermitteln. Stehen hingegen keine autoepistemischen Operatoren zur Verfügung, so stellt die Ermittlung des Maximums einer Menge ohne vorab bekannte Obermenge ein ernstes Problem dar[63]. Dies gilt gleichermaßen für den allgemeinen Fall einer Hierarchie-gesteuerten Auswahl aus einer Kandidatenmenge. Bei den erwähnten Beispielen werden die Hierarchien in Form einer Faktenmenge repräsentiert, die jedem Element eine eineindeutige Priorität zuordnet. Auf diese Weise findet eine natürliche Trennung von Regeln zur Verarbeitung und den verarbeiteten Daten statt. Soll eine Hierarchie geändert werden, so genügt die Anpassung der Prioritäten oder die Ergänzung neuer Einträge. Die von der Methode erwartete Struktur der Hierarchie wird über Constraints erzwungen. Eine Alternative wäre die Beschränkung auf eine spezielle Hierarchie, um eine Repräsentation ohne autoepistemische Operatoren zu

[63] Dass zu einem Element kein größeres Element existiert ist bereits eine nicht-monotone Aussage, die in MOBAL nur über autoepistemische Operatoren realisierbar ist.

ermöglichen. Nachteile dieses Ansatzes bestehen in der geringeren Übersichtlichkeit und der schlechteren Wartbarkeit.

Nicht-monotone Schlüsse

Der am häufigsten auftretende autoepistemische Operator im umgesetzten Modell ist \+, der bei Nichtableitbarkeit des angegebenen Literals mit dem Wahrheitswert „wahr" belegt wird. Die Verwendung dieses Operators hat verschiedene Gründe.
Zunächst sei ein Test beschrieben, der ohne vorzeitige Einschränkung der verarbeiteten Daten nur schwer ohne autoepistemische Operatoren umzusetzen ist. Es handelt sich dabei um die Negation des Test, ob ein beliebiges Medikament einer bestimmten Wirkungsgruppe gegeben wird. Auf Seite 45 ff. wird die Herleitung des Prädikates drug_effect_category/4 beschrieben, dessen Fakten die Gabe von Medikamenten gewisser Wirkungsgruppen anzeigen. Jedes verabreichte Medikament wird dosisabhängig keinem, einem oder mehreren Wirkungsgruppen zugeordnet[64]. Soll nun beispielsweise Prämisse einer Regel sein, dass zu einer Visite kein positiv inotropes Medikament gegeben wird, so kann dies nicht-monoton wie folgt ausgedrückt werden[65]:

```
\+(drug_effect_category(Patient, Time, inotrope, Drug))
```

Ohne autoepistemische Operatoren[66] ist zunächst für alle betroffenen Medikamente abzuleiten, dass sie in der aktuellen Dosierung nicht positiv inotrop wirken. Schließlich führt die Konjunktion dieser Zwischenergebnisse zum gewünschten Ergebnis. Ist die Menge der Medikamente vorab nicht bekannt, so lässt sich die Konjunktion nicht angeben. Werden hingegen nur positive Literale zu drug_effect_category/4 abgeleitet, so gibt es keine entsprechenden Einschränkungen, die Regelmenge ist kürzer, übersichtlicher und, da bei neuen Medikamenten nur Fakten zur Zuordnung zu Kategorien angegeben werden müssen, auch wartbarer. Die explizite Herleitung von Negationen erzeugt Redundanzen und kann darüber hinaus aufwendig oder unübersichtlich werden. Es ist im Einzelfall zu entscheiden, ob der Mehraufwand, die Redundanzen in der Regelmenge und die geringere Übersichtlichkeit durch die Vorteile monotoner Regeln aufgewogen werden.
Ein weiterer Grund, nicht-monotone Schlüsse in die Modellierung mit einzubeziehen ist, dass dies teilweise der menschlichen Denkweise näher kommt. Als Beispiel soll die Wahl einer Therapie dienen, falls eine Indikation, aber keine Kontraindikation vorliegt. Der Begriff „Kontraindikation" ist der medizinischen Fachsprache entnommen. Die Repräsentation solcher fachsprachlichen Begriffe vereinfacht die Diskussion mit Experten und die Modellierung des Sachbereichs. Aus diesem Grund ist eine intuitive Art der Modellierung, Kontraindikationen eigenständig zu repräsentieren und eine nicht-monotone Regel der Art

> *Wenn* Indikation *und* Kontraindikation *nicht ableitbar*,
> *dann* Wahl der Therapie.

zu formulieren.

[64] Die verwendete Abbildung auf Wirkungsgruppen wird über drugdose_to_category/4 repräsentiert.
[65] Die Variable Drug sei ungebunden, die anderen gebunden. Die Prämisse findet sich auf Seite 61 im Originalkontext.
[66] Die Verallgemeinerung von \+ ist durchaus notwendig, da der Operator count leicht dessen Funktionalität ersetzen kann. Letzterer ermittelt die Kardinalität der Menge wahrer Grundfakten, die sich durch Variablensubstitution aus einem übergebenen Ausdruck ergeben.

Grundsätzlich gilt, dass ohne nicht-monotone Schlüsse Vorteile des MOBAL-Formalismus verloren gehen. Zur Veranschaulichung sei das in Abschnitt 3.3.1 vorgestellte Beispiel zur Ableitung des Begriffs der Tachykardie in Erinnerung gerufen. Die Tabelle 3.2 zeigt, dass bestimmte Evidenzen nur mit Hilfe des Operators „\+“ zu testen sind. Sollen ohne diesen Operator anhand des Begriffs „Tachykardie“ drei Fälle unterschieden werden, „Patient tachykard“, „nicht tachykard“ und „nichts über Tachykardie bekannt“, so kann dies nicht anhand des einen Prädikats tachycardia/2 geschehen, obwohl alle drei Fällen durch eigene Evidenzen identifizierbar sind.
Durch eine explizite Repräsentation des Ereignisses „fehlender Wert“ lassen sich viele der nicht-monotonen Schlüsse umgehen. Dies führt u.a. dazu, dass in Ableitungsgraphen auf fehlenden Werten basierende Ableitungen vollständig dargestellt werden. Ein Beispiel für die Repräsentation fehlender Werte im beschriebenen Modell findet sich bei der Beschreibung des Prädikats

discrete_value/4: <patient>, <time>, <discr_attrib>, <category>.

in Abschnitt 3.6.4. Die ausgezeichnete Kategorie unknown ist ausschließlich für fehlende Werte des jeweiligen Attributs reserviert[67]. Ob Werte fehlen kann in Folge dessen wie Zugehörigkeit von Werten zu einer bestimmten Kategorie getestet werden.
Ein Vorteil der expliziten Repräsentation fehlender Werte zeigt sich auch dann, wenn die Menge der abgeleiteten Fakten in einem MOBAL-Fenster angezeigt werden. Die Information „Wert fehlt“ ist dann unmittelbar abzulesen, hingegen fehlt in der oben erwähnten Formalisierung des Begriffs „Tachykardie“ einfach das tachycardia/2-Fakt, falls die Herzfrequenz nicht verfügbar ist. Bei mehreren Prädikaten und Attributen kann schnell die Übersicht verloren gehen, welche Werte nun im einzelnen fehlen, was die Bewertung der Situation im Expertengespräch erschwert.

[67] Bei der Herleitung der Kategorie wird allerdings einmal der Operator „\+“ verwendet, da fehlende Werte auf der darunter liegenden Ebene nicht repräsentiert werden.

4. Experimente

4.1. Datensätze der Städtischen Kliniken Dortmund

Nachdem in Kapitel 3 die Umsetzung des Behandlungsprotokolls in eine operationale Form beschrieben wurde, widmet sich dieser Abschnitt der experimentellen Validierung des Modells. Dies soll, wie in Kapitel 2 bereits erläutert, unter Verwendung protokollierter Patientendaten geschehen.
Für diese Diplomarbeit standen Verlaufsdaten der Städtischen Kliniken Dortmund zur Verfügung. Es handelt sich dabei um die anonymisierten Daten von insgesamt 242 Patienten der chirurgischen Intensivstation. Die Behandlungsdauer beträgt im Schnitt 18 Tage, reicht aber von 21 Stunden bis zu 180 Tagen. Die Attribute des Datensatzes sind, soweit dies möglich war, speziell auf die vom Protokoll verwendeten Variablen zugeschnitten. Tabelle 4.1 gibt eine Übersicht der Attribute des Datensatzes. Das Zeitraster der Verlaufsdaten ist minütlich, viele der Werte sind jedoch weitaus seltener verfügbar. Insbesondere Laborwerte werden oft nur alle 4 Stunden geliefert. Neben dem Patientenzustand lassen sich aus den Daten auch die jeweils vom medizinischen Personal durchgeführten Interventionen ablesen. Ein Vergleich der aus dem Datensatz ablesbaren mit den vom Protokoll benötigten Variablen (siehe Abschnitt 3.2.2) zeigt, dass die Informationen weitestgehend, jedoch nicht vollständig zur Deckung zu bringen sind.

4.1.1. Unterschiedliche Therapien

In den Verlaufsdaten kommt eine Intervention vor, die das Protokoll nicht kennt. Dabei handelt es sich um die Gabe des Vasodilators „Nitro“. Das Protokoll verwendet gar keine Vasodilatoren. Ein direkter Vergleich der Therapien, wenn in den Verlaufsdaten die Verwendung von Nitro eingetragen ist, ist somit nicht möglich. Da die Medikation der unmittelbaren Vergangenheit und Gegenwart Teil des Patientenzustands sind, ergeben sich eventuell weitere Probleme. So gibt es im Protokoll keine Abfrage, ob in den letzten Stunden ein Vasodilator gegeben wurde, aber entsprechende Abfragen für alle dem Protokoll bekannten Medikamentengruppen, also Vasokonstringenzien, inotrope Medikamente und Diuretika.
Weniger problematisch sind die Vorschläge des Protokolls zum Medikament „Phenylephrine“, das in den Verlaufsdaten fehlt. Es kommt nur dann zum Einsatz, wenn alle anderen Vasokonstringenzien bereits in höchstmöglicher Dosierung gegeben werden. Dies ist einerseits selten in den Verlaufsdaten zu beobachten, andererseits lässt sich in diesem Fall, aufgrund der bekannten Therapieintention, dennoch ein Vergleich durchführen.
Bei Infusionen ergibt sich ein ganz anderes Problem. Während in den Datensätzen zwar die gleichen Infusionen[1] wie im Protokoll vorkommen, lässt sich die Geschwindigkeit,

[1] Treffender ist wohl „vergleichbare Infusionen“. Die Übereinstimmung liegt aber auch an einer Vorauswahl der Patientendaten. In Deutschland sind sogenannte Kolloide gebräuchlich, die im Protokoll nicht vorkommen. Dies sind Infusionen, die länger im Kreislaufsystem verbleiben. Es wurden nur

Statische Patientendaten	*Medikamente*	*Infusionen*
- Bezeichner	- Dobutamin	- Osmofundin 15%
- Gewicht	- Dopamin	- NaCl 0.9%
- Größe	- Furosemid	- Ringer-Lösung
- Geschlecht	- Nitro	- Glukose 5%
- Körperoberfläche	- Noradrenalin	- Tutofusin OPX
	- Suprarenin	

Herz-/Kreislaufwerte	*Laborwerte*	*Behandlungszeit*
- Mittlerer arterieller Druck	- Laktat	- Datum
- Systolischer Blutdruck	- Natrium	- Zeit
- Zentralvenöser Blutdruck	- Blutzucker	
- PCWP	- Hämoglobin	*Sonstige Werte*
- Herzzeitvolumen	- Creatinin	- Beatmungsdruck
- Herzfrequenz	- p_aO_2, p_vO_2	- Urinfluss
	- S_aO_2, S_vO_2	- Gesamteinfuhr
	- C_aO_2, C_vO_2	- Dialyse

Tabelle 4.1.: Attribute der verwendeten Verlaufsdaten

mit der die Infusionen liefen, nicht mehr ablesen. Das Protokoll unterscheidet deutlich zwischen schnell laufenden Infusionen (Bolus), die der Kreislaufstabilisierung dienen, und langsam laufenden Infusionen, zur Versorgung des Körpers mit Nährstoffen. In den Verlaufsdaten sind zu Infusionen nur der Zeitpunkt des Anhängens, das Volumen und die Art der Infusion verzeichnet.

4.1.2. Fehlende Patientenvariablen

Abgesehen von den teilweise unterschiedlichen Therapien ergeben sich noch weitere Probleme bei der Verwendung des Datensatzes. So lassen sich einige Patientenvariablen, die das Protokoll verarbeitet nicht aus den zur Verfügung stehenden Informationen ableiten. Während beim Herzindex eine einfache Division ausreicht und statt des rechtsatrialen auch der fast identische zentralvenöse Blutdruck verwendet werden kann, sind besonders die ohne technische Hilfsmittel bestimmbaren Variablen gar nicht verfügbar. Dazu zählen „marmorierte Beine“ und „kalte Füße“, beides Indikatoren für eine nicht hinreichende Durchblutung. Ebenso ist der Test nicht durchführbar, bei dem nach einem Druck auf die Fingerkuppe die Zeit gemessen wird, bis sich die Gefäße wieder mit Blut gefüllt haben. Flüssigkeitsansammlungen im Gewebe (Ödeme) lassen sich ebenfalls leicht optisch bestimmen, sind aber nicht im Datensatz vermerkt. Statt die entsprechenden Tests zu diesen Variablen durchzuführen, wird auf Default-Werte zurückgegriffen. Die Osmolalität des Urins wird im Protokoll alternativ über drei verschiedene Werte evaluiert, von denen jedoch keiner verfügbar ist. Statt des im Protokoll verwendeten Tests wird im weiteren Verlauf ein weniger aussagekräftiger Ersatz verwendet, der die Osmolalität anhand des Urinflusses der letzten 6 Stunden abschätzt.

Daten von Patienten verwendet, die keine Kolloide bekamen.

4.2. Vom Entscheidungshilfe- zu einem Kritiksystem

4.2.1. Anforderungen an das Modell

Die in Kapitel 3 beschriebene Umsetzung des in dieser Diplomarbeit untersuchten Behandlungsprotokolls ist so angelegt, dass für Datensätze mit geeigneten Variablen Behandlungsvorschläge abgeleitet werden können. Dabei wird der Visitenzyklus des Protokolls nachgeahmt, und es werden jeweils längerfristige Interventionsempfehlungen gegeben. Um nun die Qualität der Empfehlungen zu überprüfen sind verschiedene Ansätze denkbar. Entweder, ein Experte überprüft Vorschlag für Vorschlag, oder die im Datensatz vermerkten Therapien werden zum Vergleich herangezogen. Beim zweiten Ansatz ist die Gefahr größer, dass den zum Vergleich verwendeten Therapien Fehlentscheidungen zugrunde liegen, da sie laut Abschnitt 1.1 unter fehleranfälligen Bedingungen zustande kamen. Dieses Risiko lässt sich verringern, wenn statt dessen ein Experte in Ruhe die einzelnen Ableitungsschritte nachvollzieht. Andererseits ist das Datenvolumen sehr umfangreich, so dass eine Vorauswahl hilft, sich auf interessante Fälle zu konzentrieren. Die Aufgabe besteht daher zunächst darin, das System so zu erweitern, dass automatisch interessante Fälle gefunden werden können. Intuitiv sind das solche Fälle, bei denen der Vorschlag aus dem Protokoll nicht mit der im Datensatz vermerkten Intervention übereinstimmt. Es lässt sich annehmen, dass selten beide Therapien auf gleiche Weise falsch sind. Eine Beschränkung auf Fälle, in denen sich Protokoll und Praxis unterscheiden, bedeutet also annähernd eine Konzentration auf die Fehler. Für die gefundenen Fälle muss geklärt werden, ob wirklich ein Fehler vorliegt, da auch unterschiedliche Wege zum Ziel führen können. Dies sollte ein Experte untersuchen.
Die Aufgabe erinnert stark an Kritik-Systeme. Während das bislang beschriebene System den Entscheidungshilfesystemen zuzurechnen ist, beziehen sich Kritik-Systeme auf vom Anwender vorgegebene Lösungen. Lassen sich Schwachstellen in einer Lösung erkennen, so geben solche Systeme entsprechende Hinweise, und schlagen eventuell Verbesserungen vor. Der in [Pup99] vorgeschlagene Ansatz sieht vor, Wissensbasen von Unterstützungssystemen durch zusätzliches Kritik-Wissen zu ergänzen. In dieser Arbeit werden nur einige der vorgeschlagenen Anregungen umgesetzt. Beispielsweise würden Minimalanforderungen an Lösungen, die über das Protokoll hinaus gehen die Erhebung weiteren Wissens erfordern. Für die hier vorgestellte Anwendung genügt ein Vergleich des Protokollvorschlags und dessen Herleitung mit den wirklich durchgeführten Therapien.

4.2.2. Anpassungen des Modells

Änderung des Visitenkonzepts

Der Visitenzyklus bildet eine Art äußere Kontrollschleife für das Modell. Jede Visite bedingt, dass die relevanten Verlaufsdaten für diesen Zeitpunkt zusammengetragen werden. Nach der Bewertung des Zustands richtet sich der Zeitpunkt der nächsten Visite (siehe Abschnitt 3.6.11). Auf diese Weise geschieht die Auswahl der Visitenzeitpunkte nicht ereignisorientiert, was bei der Suche nach Auffälligkeiten ineffizient ist. Ein erster Schritt war daher eine Änderung[2], um Visiten auch diskontinuierlich, in Abhängigkeit von bestimmten Ereignissen anstoßen zu können. In diesem Fall ist die Verkettung der Visitenzeitpunkte weniger sinnvoll. Sie dient sonst der Übertragung von Wartezeiten und der

[2] Eigentlich handelt es sich um eine reine Erweiterung um einen weiteren Modus. Der Begriff „Änderung“ bezieht sich auf den Unterschied zum Visitenmodell.

Aufrechterhaltung von Notfallsituationen bis zum Abbruchkriterium. Zwischen einzelnen Visiten liegen aber teilweise mehrere Tage oder nur wenige Minuten. Der kontinuierliche Grundgedanke des Visitenzyklus lässt sich nur schlecht mit dem ereignisorientierten Modell verbinden. Änderungen der Therapievorschläge durch diese Anpassungen waren aber kaum auszumachen. Das Protokoll ist nicht auf eine Vorgänger-Visite angewiesen. Die Referenzwerte zu Variablen werden nicht über die letzte Visite bezogen, sondern unmittelbar durch eine Datenbankanfrage, so dass die Bewertung der Werteveränderungen von den Anpassungen nicht beeinträchtigt wird.
Als Ereignisse, um Visiten anzustoßen bieten sich im Datensatz vermerkte Interventionen an. Dieses Vorgehen wird auch in [MIB+00] erfolgreich angewandt. Im Moment einer Intervention ist zu erwarten, dass die Patientenvariablen die Entscheidung für die Intervention am Besten stützen. Auch falls Visiten in der Klinik Änderungen der Medikation zu bestimmten Zeitpunkten begünstigen, so ist in diesen Momenten die größte Abhängigkeit zwischen beobachteter Intervention und Patientenzustand zu erwarten. Zu anderen Zeitpunkten ist unklar, ob vielleicht nur deshalb keine Änderungen erfolgen, weil kein Arzt am Krankenbett ist. Indem nur Zeitpunkte ausgewertet werden, zu denen die Interventionen gut durch die Daten gestützt werden, steigt effektiv die Qualität der Daten. Diese Methode eignet sich allerdings nur zur Validierung des Behandlungsprotokolls, und nicht zur Validierung der Daten[3]. Bei letzterer wäre unbedingt auf versäumte Interventionen aufmerksam zu machen.
Die „Fälle“ aus Abschnitt 4.2.1 können als Patientenzustände inklusive der im Datensatz verzeichneten Interventionen aufgefasst werden. Patientenzustände umfassen in diesem Sinne auch Messwerte und Ereignisse aus der Vergangenheit, die in den Prozess der Therapie-Findung mit eingehen. Referenzwerte für die Verlaufsbewertung und Ereignisse wie „Innerhalb der letzten Stunde wurde ein Diuretikum gegeben.“ stellen Beispiele dar. Die Ereignisse, die Visiten anstoßen, werden zunächst als Fakten repräsentiert. Für Änderungen der Medikation geschieht dies anhand des Prädikats

before_change/5: <patient>, <time>, <drug>, <dose>, <drug_change_dir>,

für angehängte Infusionen anhand von

before_infusion/4: <patient>, <time>, <infusion>, <infusion_ml>.

Diese Prädikate bilden die Schnittstelle der Kritik-Ebene zu den Interventionen des Datensatzes. Der in den Fakten angegebene Zeitpunkt liegt jeweils eine Minute *vor* der Intervention. Beim ersten Prädikat dienen die letzten beiden Argumentstellen der Angabe der neue Dosis und der Änderungsrichtung. Beim zweiten Prädikat werden jeweils die Infusion und das angehängte Volumen spezifiziert.
Fakten dieser Prädikate stoßen nicht nur Visiten an, sondern darüber hinaus auch einen Vergleich des abgeleiteten Therapievorschlags mit dem Ereignis im Datensatz. Es wird nun zunächst näher auf mögliche Bewertungsgrundlagen für Vergleiche eingegangen.

Vorüberlegungen zu Vergleichen

Für die Kritik eines Falles stehen unterschiedliche Abstraktionsebenen zur Verfügung. Eine sehr einfache Methode wäre ein Vergleich der konkreten Interventionen mit dem Vorschlag des Protokolls. Beispielsweise könnte eine Erhöhung von Adrenalin von 0.04

[3]Eine alternative Aufgabenstellung wäre, in protokollierten Daten nach medizinischen Fehlern zu suchen (vgl. Abschnitt 1.1). Dies erfordert allerdings bereits ein gut ausgearbeitetes Protokoll.

μg/kg/min auf 0.06 μg/kg/min in den Datensätzen verzeichnet sein, während das Protokoll vorschlägt, Dobutamin von 15 μg/kg/min auf 20 μg/kg/min zu erhöhen. Auf der untersten Ebene lässt sich nun lediglich festhalten, dass die Interventionen sich unterscheiden, da sie sich auf verschiedene Medikamente beziehen. Auf den zweiten Blick handelt es sich jedoch in beiden Fällen um positiv inotrope Medikamente. Im Inferenzprozess des Systems wird zunächst die Erhöhung eines Medikaments dieser Medikamenten*gruppe* abgeleitet, bevor dann ein konkretes Medikament gewählt wird. Indem im Protokoll die Reihenfolge von Adrenalin und Dobutamin in der Hierarchie inotroper Medikamente geändert wird, leitet das Protokoll die Intervention des Datensatzes ab[4]. Dieses Beispiel zeigt, dass sich neben den konkreten Interventionen auch die abstrakteren Ebenen der Lösungshierarchie für einen Vergleich eignen. Wenn sich bereits auf der Ebene abstrakter Therapien Protokoll und Praxis unterscheiden, dann ist diese Ebene für die Kritik interessanter. Stimmen die groben Therapien überein, so kann nach feineren Differenzen gesucht werden, etwa beim Medikament oder sogar der vorgeschlagenen Dosis.

Abstraktion der Interventionen aus dem Datensatz

Um Interventionen auf verschieden abstrakten Ebenen vergleichen zu können, muss von den Dosierungen der Medikamente und angehängten Infusionsvolumen im Datensatzes auf die zugrunde liegenden intendierten Therapien rückgeschlossen werden. Zugrunde liegende Therapien sind dabei etwa im Sinne der in den Abschnitten 3.6.6 und 3.6.7 vorgestellten Abstraktionen[5] zu verstehen.

Für die Zuordnung von Medikamenten zu den im Protokoll definierten Gruppen kann auf das in Tabelle 3.1 zusammengefasste Wissen zurückgegriffen werden:

$$
\begin{array}{ll}
 & \mathsf{before\_change(Patient,\ Time,\ Drug,\ Dose,\ Direction)} \\
\wedge & \mathsf{drugdose\_to\_category(Drug,\ LowerBound,\ UpperBound,\ Category)} \\
\wedge & \mathsf{Dose} \geq \mathsf{LowerBound} \\
\wedge & \mathsf{Dose} \leq \mathsf{UpperBound} \\
\longrightarrow & \mathsf{performed\_therapy(Patient,\ Time,\ Drug,\ Category,\ Direction).}
\end{array}
$$

Die Abstraktion der Interventionen im Datensatz entspricht auf natürliche Weise einer Umkehrung der Verfeinerung von Therapien im Protokoll. Die schließlich abgeleiteten Beschreibungen von Interventionen sind von der Form „Dosis eines Vasokonstriktors erhöht“ etc., analog zu den Beschreibungen anhand des Prädikats therapy/4.

Schwieriger gestaltet sich die Abstraktion bei Infusionen. Wie im Abschnitt 4.2.2 bereits erwähnt, fehlt im Datensatz die Information, wie schnell die angehängten Infusionen laufen. Die abstrakteren Formen der Infusionen im Protokoll unterscheiden sich jedoch besonders durch die unterschiedliche Durchlaufgeschwindigkeit, weniger durch die Zusammensetzung. Um auch an dieser Stelle einen Vergleich auf abstrakterer Ebene durchführen zu können, wurden mit Expertenhilfe Heuristiken formuliert, die die als Bolus verabreichten Infusionen von den übrigen unterscheiden. Hierzu wurden spezielle, teils technisch bedingte Eigenschaften des Datensatzes ausgenutzt. Infusionen wird der jeweilige heuristisch zugeordnete Modus über Fakten zum Prädikat

[4] Die Hierarchie findet sich auf Seite 62. Adrenalin entspricht dort supra, Dobutamin entspricht dobutrex.

[5] Gemeint sind Abstraktionen, wie sie im Rahmen der Prädikate therapy_candidate/4 und therapy/4 auftreten.

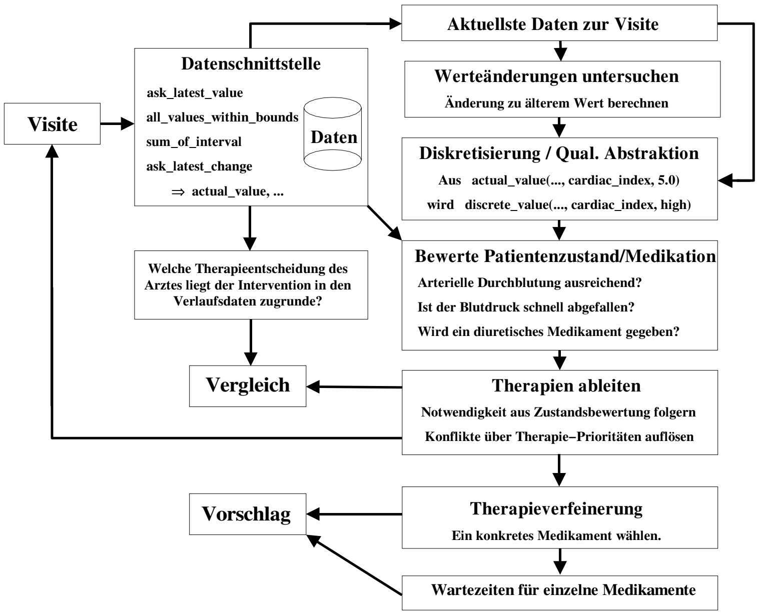

Abbildung 4.1.: Architektur des Kritik-Systems

performed_infusion/5: <patient>, <time>, <infusion>, <infusion_ml>,
<perf_infu_mode>.

zugeordnet. Die Sorte <perf_infu_mode> enthält die Konstanten bolus, für Bolusgaben, normal, für alle übrigen Infusionen und unknown, falls keine Zuordnung möglich ist.

Bewertungskriterien

Für den Vergleich zwischen der Therapie aus dem Datensatz und der Protokollempfehlung sind verschiedene Maße denkbar. Ziel ist eine übersichtliche Erfassung, die interessante Fälle erkennen lässt. Welches Maß zu einer geeigneten Partitionierung der Menge aller Fälle führt hängt auch vom Zusammenspiel zwischen Protokoll und Daten ab. Bei sehr wenigen Abweichungen wird es etwa interessant sein, die „Fast-Treffer“ besser zu unterteilen. Beim umgesetzten Protokoll und dem verwendeten Datensatz schienen erste Betrachtungen jedoch auf wenige Übereinstimmungen hinzudeuten. Entsprechend schien es angemessener, die Vergleiche auf die abstrakteren Ebenen zu konzentrieren. Die Abbildung 4.1 gibt die entsprechend angepasste Architektur des Modells wieder.
Um die Bewertung verschiedener Fälle übersichtlich darstellen zu können, wurde der Signatur ein eigenes Prädikat

critiquing_status/5: <patient>, <time>, <critiq_statement>,
<therapy_object>, <therapy_mode>.

hinzugefügt. Grundlage der Vergleiche sind einerseits Fakten zu den oben vorgestellten Prädikaten performed_infusion/5 und performed_therapy/5, die die intendierten Therapi-

en zu den im Datensatz vermerkten Interventionen wiedergeben, andererseits akzeptierte Therapien (therapy/4). Es zeigte sich nach ersten Versuchen Bedarf für eine noch feinere Unterteilung der Nicht-Übereinstimmungen, deshalb wurden außerdem noch die vom Protokoll erzeugten Therapiekandidaten (therapy_candidate/4) in das Bewertungsmaß aufgenommen.
Es folgen die Regeln zu den unterschiedlichen gebildeten Kategorien[6], repräsentiert durch unterschiedliche Konstanten der Sorte <critiq_statement>. Zunächst werden die medikamentösen Therapien verglichen:
Die Therapien stimmen auf abstrakter Ebene überein (same_therapy):

performed_therapy(Pat, T, Drug, DrugCateg, Direction)
∧ therapy(Pat, T, DrugCateg, Direction)
⟶ critiquing_status(Patient, T, same_therapy, DrugCateg, Direction).

Die abstrakte Intervention des Datensatzes befindet sich nicht einmal unter den Therapiekandidaten des Protokolls (performed_is_no_candidate):

performed_therapy(Pat, T, Drug, Obj, Mode)
∧ \+(therapy_candidate(Pat, T, Obj, Mode))
⟶ critiquing_status(Pat, T, performed_is_no_candidate, Obj, Mode).

Das Protokoll empfiehlt für die gleiche Medikamentengruppe die gegenteilige Änderung, also eine Erhöhung, wenn die Dosis im Datensatz herabgesetzt wird und umgekehrt (clearly_contradicting):

performed_therapy(P, T, Drug, DrugCateg, PerfDir)
∧ therapy(P, T, DrugCateg, ProposedDir)
∧ contradicting_direction(PerfDir, ProposedDir)
⟶ critiquing_status(P, T, clearly_contradicting, DrugCateg, PerfDir).

Dabei sind „entgegengesetzte Richtungen" anhand von contradicting_direction/2 definiert als ein decrease und increase oder decrease und urgent_increase.
Der Vergleich zwischen durchgeführter und vorgeschlagener Infusionstherapie geschieht gemäß folgender Regeln:

performed_infusion(Patient, Time, Infu, Volume, PerfMode)
∧ therapy(Patient, Time, infusion, PropMode)
∧ infusion_critiq(PerfMode, PropMode, Statement)
⟶ critiquing_status(Patient, Time, Statement, PerfMode, PropMode).

performed_infusion(Patient, Time, Infu, Volume, PerfMode)
∧ \+(therapy(Patient, Time, infusion, PropMode))
∧ infusion_critiq(PerfMode, none, Statement)
⟶ critiquing_status(Patient, Time, Statement, PerfMode, none).

[6] Wie schon angedeutet sind hier noch zahlreiche andere Kategorien denkbar, die teilweise auch verwendet wurden. Hier wird eine Variante vorgestellt, die die betrachteten Fälle gut unterteilt.

Durchgeführt	Vorgeschlagen	Bewertung
bolus	bolus	same_infusion
bolus	maintenance	different_infusion
bolus	keep_open	different_infusion
bolus	none	different_infusion
normal	bolus	different_infusion
normal	maintenance	same_infusion
normal	none	different_infusion

Tabelle 4.2.: Bewertung von Paaren durchgeführter und vorgeschlagener Infusionstherapien

Anhand des Prädikats infusion_critiq/3 wird möglichen Paaren jeweils eine Bewertung zugeordnet. Unklare Fälle bleiben dabei unberücksichtigt. Tabelle 4.2 gibt die Zuordnung wieder.

4.2.3. Experimente mit dem Kritik-System

Durch die beschriebene Kritik-Ebene lassen sich die Bewertungen für die Therapievorschläge übersichtlich darstellen. Abbildung 4.2 zeigt beispielhaft einen Auszug der Bewertungen zu Patient 6983.
Aufgrund des Datenvolumens wurde im Rahmen der Experimente bislang keine umfassende Untersuchung aller Datensätze vorgenommen. Die Auswahl der betrachteten Verlaufsdaten geschah jedoch nicht willkürlich, sondern mit Hilfe einiger Datenbank-Anfragen, die dazu dienten, möglichst unterschiedliche Datensätze mit bestimmten Eigenschaften zu finden. Zu den Auswahl-Merkmalen zählten etwa das Auftreten hypotoner Krisen, also eines sehr niedrigen arteriellen Mitteldrucks, und die Gabe bestimmter Medikamente. Da das Protokoll keine Vasodilatoren kennt, wurden überwiegend Behandlungsverläufe untersucht, bei denen solche Medikamente nicht zum Einsatz kamen.
Bei ersten Experimenten wurden die Daten von ca. 10 Patienten ausgewertet. Es fanden sich nur sehr wenige Übereinstimmungen zwischen den Vorschlägen des Protokolls und den protokollierten Interventionen. Der Anteil der Übereinstimmungen (same_therapy) lag bei keinem Datensatz über 5 Prozent. Einige im Laufe der Zeit gewonnene Erkenntnisse halfen dabei, die Resultate zu verbessern. Dies bezieht sich insbesondere auf eine Verbesserung der Datenqualität, etwa durch Plausibilitätstests, um Messfehler zu erkennen und herauszufiltern. Hierauf wird in Abschnitt 4.3.2 nochmal eingegangen. Weiterhin gelang es einige Ungenauigkeiten bei der Interpretation der Daten zu identifizieren. Beispielsweise wurden bei den intrathorakalen Blutdrücken zunächst Verfälschungen durch künstliche Beatmung außer Acht gelassen, einige Variablen wurden zu lange als noch aktuell eingestuft[7] und ein Ersatztest, der fehlende Attribute kompensieren sollten, erwies sich als unangemessen und mussten überarbeitet werden. Obwohl sich die Ergebnisse im Laufe der Zeit signifikant verbessert haben, weichen die Vorschläge für einige Patienten noch deutlich von den protokollierten Interventionen ab.
Tabelle 4.3 fasst für das beschriebene Experiment die Ergebnisse des letzten Laufes zusammen. Bei den medikamentös ausgerichteten Therapien sind besonders die Übereinstimmungen (same_therapy) und entgegengesetzten Therapien (clearly_contradicting)

[7] An dieser Stelle musste das Protokoll ergänzt werden.

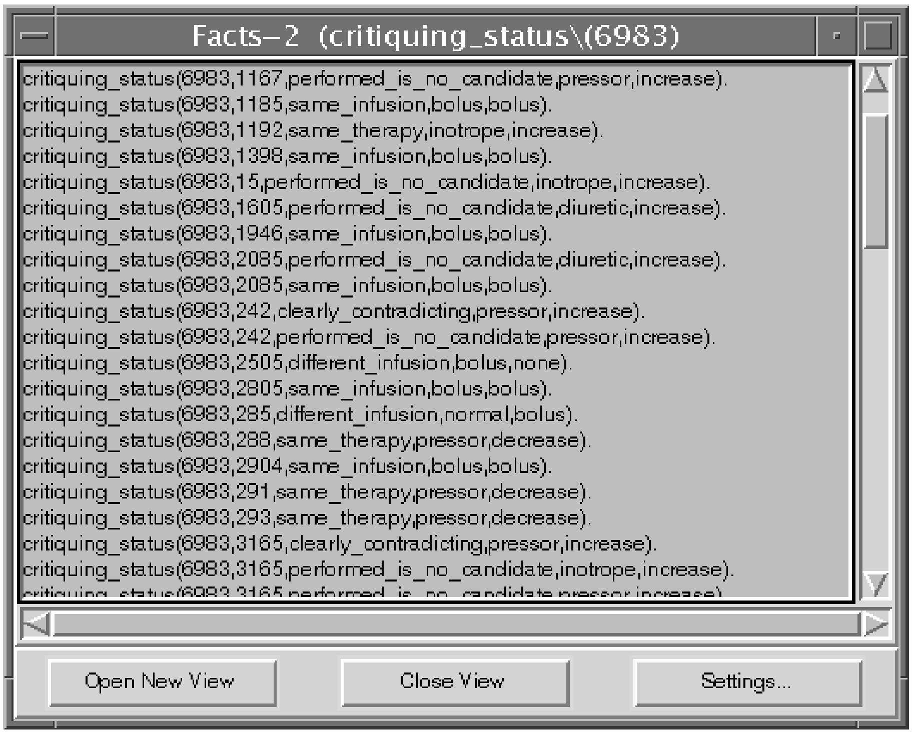

Abbildung 4.2.: Auszug der Kritik-Ergebnisse für Patient 6983

aussagekräftig. Befindet sich die durchgeführte Therapie nicht unter den Therapiekandidaten (performed_is_no_candidate), so kann dies hingegen unterschiedliche Gründe haben, etwa dass das Protokoll zum Erreichen bestimmter Ziele andere Mittel vorsieht. Die Bewertungen der medikamentösen Therapien zeigen sich Patienten-abhängig uneinheitlich. Bei Patient 7023 überwiegen die Übereinstimmungen, bei Patient 5826 die klaren Widersprüche. Insgesamt sind besonders die klaren Widersprüche erklärungsbedürftig. Bei den Infusionstherapien zeigen sich durchgängig Unterschiede zwischen Protokoll und Daten, die sich in dieser Ausprägung nicht mehr auf die lediglich heuristische Zuordnung der Infusionstypen zurückführen lassen.
Es gelang nach Treffen mit einem Experten einige medizinische Gründe für abweichende Therapievorschläge zusammenzutragen. Dabei wurde, ausgehend von problematischen Protokollvorschlägen, besonders auf die in Abschnitt 3.7 beschriebenen Methoden zurückgegriffen. Die medizinischen Erkenntnisse sollen aufgrund der anderen Zielsetzung dieser Arbeit nur knapp dargestellt werden.
Das Protokoll weicht in folgender Hinsicht von der in Dortmund gängigen Praxis ab:

- Das Protokoll verwendet Diuretika statt blutdrucksenkende Mittel bei Hypertonie. Blutdrucksenkende Mittel kommen im Protokoll nicht vor.

- Die Zielwerte für diverse Variablen weichen im Protokoll von den in Deutschland üblichen ab. So wird in den in den Datensätzen beobachtbaren Interventionen beispielsweise ein arterieller Mitteldruck zwischen 90 und 100 mmHg angestrebt, im Protokoll wird bereits 60 mmHg als ausreichend betrachtet. Entsprechend schlägt das Protokoll bei Werten zwischen den angestrebten Grenzen eine Verminderung

Patient	Medikamentöse Therapie			Infusionstherapie	
	gleich	verschieden	entgegen	gleich	verschieden
4848	17	24	11	13	35
4900	0	0	0	14	4
4910	1	2	1	14	15
4965	16	14	27	2	82
4974	1	0	0	0	25
5039	0	0	0	24	113
5105	9	17	9	3	160
5150	2	1	0	1	35
5403	4	2	1	4	29
5438	4	1	1	4	34
5588	9	7	5	1	78
5641	4	0	0	4	56
5765	2	0	0	1	21
5820	4	0	0	8	14
5826	7	28	16	3	12
6036	4	2	1	0	22
6043	1	3	1	5	17
6065	9	7	4	1	32
6201	2	16	1	3	15
6350	8	13	11	9	23
6425	4	1	2	2	37
6430	7	4	1	2	22
6262	7	9	5	15	17
6479	25	22	11	7	26
6553	13	14	6	1	55
6611	26	24	17	4	40
6983	15	22	9	13	11
6986	16	21	12	9	13
6994	1	5	2	1	17
7019	0	3	0	13	22
7023	6	2	1	1	32

Zuordnung zu Konstanten:

Medikamentöse Therapie		Infusionstherapie	
gleich :	same_therapy	gleich :	same_infusion
weicht ab :	performed_is_no_candidate	weicht ab :	different_infusion
entgegen :	clearly_contradicting		

Tabelle 4.3.: Zusammenfassung der Kritik-Ergebnisse

von Vasokonstringenzien vor, während diese im Datensatz oft sogar noch erhöht werden.

- Eine typische Behandlungsstrategie in Deutschland für bestimmte Patientengruppen ist der sogenannte „hyperdyname" Ansatz. Hierbei wird neben einem blutdrucksenkenden Medikament (Vasodilator) noch ein positiv inotropes Medikament verabreicht. Dieser Aspekt fehlt ganz im Protokoll.
- Das Protokoll verwendet keine „Kolloide" als Infusionen. Dabei handelt es sich um Flüssigkeiten, die länger im Kreislauf verbleiben als die vom Protokoll ausschließlich verwendeten „Kristalloide". Für die Experimente wurde dies durch eine Beschränkung auf Datensätze kompensiert, in denen keine Kolloide vorkommen.
- Die Ausrichtung des Protokolls ist überwiegend blut*druck*orientiert. In Deutschland wird hingegen mehr auf eine hinreichende Durchblutung (Blutzirkulation) geachtet.

4.3. Test der protokollinternen Konsistenz

4.3.1. Beschreibung eines weiteren Experiments

Im letzten Abschnitt wurde das Protokoll durch Vergleiche der Vorschläge mit tatsächlich in den Städtischen Kliniken Dortmund durchgeführten Interventionen experimentell validiert. Darüber hinaus lassen sich noch einige weitere Aspekte des Protokolls validieren. Bei näherer Betrachtung erscheint der Aufwand des Protokolls hinsichtlich zeitlicher Vorverarbeitung und der Komposition mehrerer redundanter Variablen zu einem Gesamtwert recht gering. Im Abschnitt 1.5.2 werden mehrere Arbeiten zur Vorverarbeitung intensivmedizischer Zeitreihen vorgestellt. Besonderes Gewicht bei diesen Methoden hat Robustheit gegenüber Rauschen und Artefakterkennung. Im Gegensatz dazu finden sich im verwendeten Protokoll keinerlei derartige Verfahren. Variablen werden bestenfalls hierarchisch nach Verfügbarkeit ausgewählt, ohne Redundanzen auszunutzen. Die Diskretisierung von Variablen geschieht anhand des Messwertes zu einer festen Minute durch Zuordnung zu einem durch absolute Werte beschriebenen Intervall.
Diese Beobachtungen legen ein Experiment nahe, das die Stetigkeit der Protokollvorschläge über kürzere Zeitabschnitte überprüft. Die Idee dabei ist, dass ohne Artefakterkennung und Glättung der Zeitreihen bei sehr einfachen Verfahren zur Diskretisierung usw., die Auswirkungen von Rauschen, zufälligen Schwankungen der echten Werte und Artefakten zu häufigen Änderungen der vorgeschlagenen Therapie führen könnten. Um dies zu überprüfen genügt es, Visiten im Minutentakt anzustoßen. Wegen des längerfristigen Charakters der Entscheidungen bei Visiten ist eine zu hohe Änderungsfrequenz ein schlechtes Zeichen. Darüber hinaus können kontinuierlich erzeugte Therapievorschläge zu übersichtlicheren zeitlichen Intervallen zusammengefasst werden. Auf diese Weise kann der Verlauf der Protokollvorschläge besser mit dem tatsächlich durchgeführten Therapieverlauf verglichen werden[8].
Die Umgebung MOBAL eignet sich nicht für das beschriebene Experiment. Einerseits werden zu viele Fakten erzeugt, wenn zu jeder Minute des Datensatzes eines Patienten Vorschläge abgeleitet werden sollen. Andererseits beliefe sich die Rechenzeit bei im Schnitt 25000 erfassten Minuten pro Patient wohl jeweils auf mehrere Wochen. Um das

[8] Dabei ist allerdings zu beachten, dass die Intervalle nur annähernd das Protokollverhalten wiedergeben. Für das Protokoll ist ausdrücklich der nicht kontinuierliche Visitenzyklus vorgeschrieben.

Experiment dennoch durchführen zu können, war die Programmierung eines schlanken Interpreters für das Modell geboten. In der Programmiersprache PROLOG[9] ließen sich mit einfachen Mitteln die MOBAL-Regeln parsen und über Rückwärtsinferenz Therapien herleiten. Die vierwertige Logik von MOBAL konnte über eine Aufteilung jedes MOBAL-Prädikates in jeweils zwei PROLOG-Prädikate nachgebildet werden. Die Datenschnittstelle war über einfache Dateioperationen zu realisieren. Da jede Minute ausgewertet wird, kann der Interpreter die Verlaufsdaten sequentiell durchlaufen. Dabei wurden die Daten des jeweils letzten Tages (bezogen auf den aktuell ausgewerteten Zeitpunkt) in einem Cache verwaltet, um auch Anfragen bezüglich älterer Werte und Ereignisse verarbeiten zu können[10]. Da der Interpreter mit nicht beschränkter Rückwärtsinferenz arbeitet, waren allerdings einige Anpassungen notwendig, um Zyklen in der Inferenz abzufangen, die ansonsten in endloser Rekursion enden würden. Dies bezieht sich besonders auf den Visiten*zyklus*.

Im Gegensatz zum Experiment mit dem Kritiksystem ist das hier beschriebene Experiment mit dem Visitenzyklus konform, obwohl der Auswertungstakt minütlich ist. Wird in der Visite zum Zeitpunkt t die nächste Visite für den Zeitpunkt $t + 30$ festgesetzt, so wird zum Zeitpunkt $t + 30$ der Zeitpunkt t als Vorgänger verwendet. Wartezeiten werden etwa auf diese Weise übertragen. Eine Visite kann bei der verwendeten Umsetzung des Protokolls nur dann mehrere Vorgänger haben, wenn einer der Vorgänger eine nicht beendete Notfallsituation darstellt. In diesem Fall wird nur der Notfall fortgeführt und die anderen Vorgänger werden ignoriert.

Das Inferenzverhalten des Systems entspricht also dem von MOBAL interpretierten Modell mit Visitenzyklus. Die inferierten Fakten bleiben allerdings nicht länger im Speicher, als sie Auswirkungen auf die weiteren Berechnungen haben können. Eine Verwendung spezifischerer (stärkerer) Problemlösungsmethoden wäre aufgrund der Struktur des Wissens möglich gewesen, die Effizienz war jedoch auch mit einfacher Rückwärtsinferenz hinreichend. Die Gefahr das Systemsverhalten gegenüber MOBAL zu verändern blieb auf diese Weise geringer.

Die Ergebnisse werden, wie oben motiviert, in eine kompaktere Form überführt, indem aufeinander folgende gleiche Therapien, Interventionsvorschläge, usw. zu Intervallen zusammengefasst werden.

4.3.2. Ergebnisse zur protokollinternen Konsistenz

Das beschriebene Experiment wurde für die Daten verschiedener Patienten durchgeführt. Vollständig minütlich durchlaufen wurden die Behandlungsdaten von insgesamt 10 Patienten. Aus der Fülle der Fakten, die sich jeweils auf ganze Zeitintervalle beziehen, werden im folgenden nur die Fakten zu akzeptierten Therapien aufgelistet. Beim entsprechenden Prädikat **therapy/4** bezeichnet die zweite Stelle nun halboffene Zeitintervalle der Art $[t_1, t_2)$ statt Zeitpunkte. Die Einheit für die Intervallgrenzen sind wieder die Minuten seit Beginn der Datenerfassung für den Patienten.

Die Therapieintervalle bieten einen recht guten Überblick, wie die Protokollvorschläge im Laufe der Behandlung in etwa[11] aussähen. Die Tabellen 4.4 und 4.5 zeigen zwei Beispiele für unstetige Behandlungsverläufe, wie sie in unterschiedlicher Ausprägung in den

[9] Für eine Einführung in PROLOG sei auf [Bra87] verwiesen.

[10] Ereignisse und Werte, die mehr als 24 Stunden zurückliegen, sind für das Protokoll nicht mehr relevant.

[11] Das „in etwa“ bezieht sich auf die Tatsache, dass das Behandlungsprotokoll weitaus seltener angestoßen wird.

	therapy(4999,	[19815,	20727],	pressor,	decrease)
×	therapy(4999,	[20725,	20727],	infusion,	maintenance)
×	therapy(4999,	[20727,	20728],	pressor,	increase)
×	therapy(4999,	[20727,	20728],	infusion,	bolus)
	therapy(4999,	[20728,	21085],	infusion,	maintenance)
	therapy(4999,	[20728,	21346],	pressor,	decrease)
×	therapy(4999,	[21346,	21347],	pressor,	increase)
	therapy(4999,	[21347,	21482],	pressor,	decrease)
×	therapy(4999,	[21482,	21483],	pressor,	increase)
×	therapy(4999,	[21483,	21484],	pressor,	decrease)
×	therapy(4999,	[21484,	21485],	pressor,	increase)
	therapy(4999,	[21485,	21497],	pressor,	decrease)
×	therapy(4999,	[21497,	21498],	pressor,	increase)
	therapy(4999,	[21085,	21565],	infusion,	bolus)
	therapy(4999,	[21565,	21805],	infusion,	maintenance)

Tabelle 4.4.: Ausschnitt aus den minütlichen Therapien in Intervallform für Patient 4999. Für die betreffenden Zeit ist keine Intervention im Datensatz vermerkt.

Ergebnissen zu allen der 10 vollständig minütlich durchlaufenen Datensätze zu finden waren. Da sich intuitiv die Therapieentscheidungen des Protokolls zumindest kurzfristig selten ändern sollten, sind in den Tabellen auffällig kurze Intervalle (maximal zwei Minuten) gekennzeichnet. Zu Tabelle 4.5 sind die medikamentösen Interventionen des Behandlungsprotokolls für den betreffenden Zeitabschnitt aufgeführt. Zum Zwecke der Vergleichbarkeit sind auch die jeweiligen Wirkungsgruppen vermerkt[12].

Diese Ergebnisse halfen in unterschiedlicher Weise bei der Analyse des Protokolls. Die erste Anwendung war eine Untersuchung der unstetigen Verläufe, indem gezielt zu den entsprechenden Zeitpunkten Visiten angestoßen wurden. Ein Experte auf dem Gebiet der Intensivmedizin stellte daraufhin unmittelbar fest, dass die verwendeten Verlaufsdaten unvalidiert waren. Verlaufsdaten werden als unvalidiert bezeichnet, wenn sie unmittelbar von den Messgeräten stammen. Verschiedene Fehlerquellen, wie das Berühren von Messkontakten oder sich bewegende Patienten führen teilweise zu kurzen Artefakten, die sich in unvalidierten Verlaufsdaten wiederfinden. Es wurde eine einfache Artefaktbereinigung durchgeführt, die auf einem Plausibilitätstest beruht. Beispielsweise kann aus medizinischen Gründen ausgeschlossen werden, dass sich ein arterieller Blutdruck innerhalb einer Minute um 25 mmHg ändert. Tritt eine solche Schwankung auf, so wird der fragliche Wert für ungültig erklärt. Die Artefaktbereinigung und einige weitere Korrekturen führten zu einer höheren Datenqualität und dadurch zu besseren Kritik-Ergebnissen (siehe Abschnitt 4.2.3). Die übersichtliche Darstellung in Intervallform half also dabei, Missstände zu bemerken und deren Ursache zu ermitteln. Aufgrund der etwas anderen Zielsetzung dieser Arbeit wurden die Auswirkungen einzelner Korrekturen der Ausgangsdaten und Ergänzungen nicht derart umfangreich untersucht, dass statistisch fundierte Aussagen möglich wären. Eine gezielte Untersuchung der Auswirkungen verschiedener Vorverarbeitungsoperatoren steht noch aus, und kann Bestandteil nachfolgender Arbei-

[12] Dabei entsprechen sich „pressor" und „Vasokonstriktor", sowie „inotrope" und „positiv inotrop". Keines der Medikamente wirkt diuretisch.

	therapy(6983,	[46,	827],	infusion,	bolus)
	therapy(6983,	[230,	706],	pressor,	decrease)
×	therapy(6983,	[526,	527],	diuretic,	decrease)
	therapy(6983,	[707,	1075],	pressor,	decrease)
×	therapy(6983,	[804,	805],	diuretic,	decrease)
	therapy(6983,	[827,	847],	infusion,	keep_open)
	therapy(6983,	[847,	1003],	infusion,	bolus)
×	therapy(6983,	[1003,	1005],	infusion,	keep_open)
×	therapy(6983,	[1005,	1006],	infusion,	bolus)
	therapy(6983,	[1006,	1066],	infusion,	keep_open)
	therapy(6983,	[1011,	1066],	diuretic,	increase)
×	therapy(6983,	[1075,	1076],	pressor,	increase)
×	therapy(6983,	[1075,	1076],	infusion,	bolus)
×	therapy(6983,	[1076,	1077],	pressor,	decrease)
	therapy(6983,	[1077,	1081],	pressor,	increase)
	therapy(6983,	[1077,	1081],	infusion,	bolus)
×	therapy(6983,	[1081,	1083],	pressor,	decrease)
×	therapy(6983,	[1083,	1084],	pressor,	increase)
×	therapy(6983,	[1083,	1084],	infusion,	bolus)
×	therapy(6983,	[1084,	1085],	pressor,	decrease)
×	therapy(6983,	[1085,	1087],	pressor,	increase)
×	therapy(6983,	[1085,	1087],	infusion,	bolus)
	therapy(6983,	[1087,	1090],	pressor,	decrease)

Minute	*Medikament*	*Neue Dosis*		*Vasokonstr.*	*Pos. inotrop*
40	arteren	0.70	↑	×	
45	supra	0.58	↑	×	
242	supra	0.70	↑	×	
242	dopamin	2.92	↑		
288	arteren	0.58	↓	×	
291	supra	0.58	↓	×	
293	supra	0.53	↓	×	
345	supra	0.41	↓	×	
356	arteren	0.41	↓	×	
362	supra	0.29	↓	×	
502	supra	0.23	↓	×	×
506	dobutrex	20.47	↓		×
507	supra	0.12	↓	×	×
508	supra	0.06	↓	×	×
574	dobutrex	17.54	↓		×
518	supra	0.0	↓	×	×
525	arteren	0.29	↓	×	
601	arteren	0.23	↓	×	
801	dobutrex	14.62	↓		×
825	arteren	0.29	↑	×	
888	arteren	0.35	↑	×	

Tabelle 4.5.: Ausschnitt aus den minütlichen Therapien in Intervallform für Patient 6983, zusammen mit den im Datensatz vermerkten medikamentösen Interventionen.

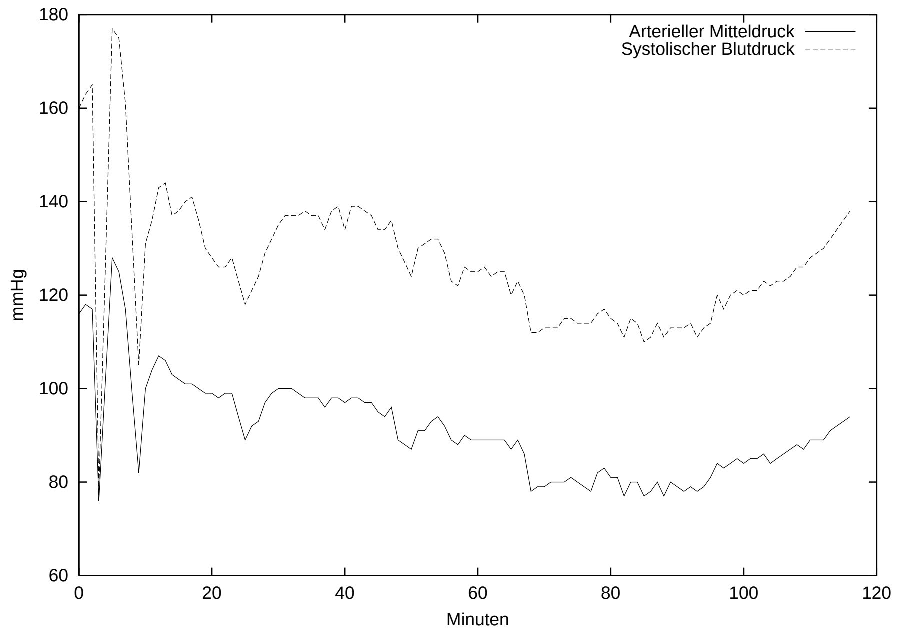

Abbildung 4.3.: Exemplarischer Verlauf eines Blutdrucks

ten sein.

Anhand der erzeugten Therapieintervalle gelang es weiterhin zu zeigen, dass Vergleiche von Attributen mit festen Grenzwerten im praktischen Einsatz systematisch unstetige Ergebnisse produzieren können. Dies gilt in besonderem Maße für den mittleren arteriellen Blutdruck und die Grenze dieses Wertes, ab der von einer bedrohlichen Hypotonie gesprochen werden kann. Schwankt der Blutdruck um den Vergleichswert, so wird „abwechselnd“ eine nicht lebensbedrohliche[13], aber schwere Hypotonie diagnostiziert, die mit dem Heraufsetzen von Vasokonstringenzien bekämpft wird, und eine leichtere Hypotonie, die zum Herabsetzen von Vasokonstringenzien führt. Dieses widersprüchlich erscheinende Verhalten ist nicht auf Artefakte zurückzuführen, sondern auf die natürlicherweise recht starken Schwankungen des arteriellen Mitteldrucks. Abbildung 4.3 zeigt exemplarisch den Verlauf dieser Variablen über einen Zeitraum von 2 Stunden. Dementsprechend gelang es mit den bislang durchgeführten Maßnahmen nicht, die auffälligen Schwankungen zwischen den Therapien (pressor, increase) und (pressor, decrease) in den Tabellen zu beseitigen. Es ist bereits abzusehen, dass auch Glättungsverfahren nur begrenzt das häufige Passieren solch fester Vergleichswerte einzuschränken vermögen.

Bei der Bewertung dieser Ergebnisse darf nicht außer Acht gelassen werden, dass das Protokoll üblicherweise nicht minütlich angestoßen wird. In der oben beschriebenen Situation dürften maximal 30 Minuten zwischen zwei aufeinanderfolgenden Visiten liegen. Dass eine minimale Veränderung des Patientenzustands bzw. eine minimale Verschiebung des Visitenzeitpunkts für eine längere Zeit zum gegenteiligen Therapievorschlag führt, ist intuitiv jedoch ebenfalls problematisch. Speziell im oben beschriebenen Beispiel werden oszillierende Änderungen der Dosis durch den Mechanismus der Wartezeiten abgefangen.

[13] Kommen weitere Kriterien hinzu, so wird evtl. eine lebensbedrohliche Hypotonie diagnostiziert, die zu einem Notfall führt. Das Notfallprotokoll hat eine explizite Abbruchbedingung, was das Problem verringert.

Die zugrundeliegende Art von Vergleich ist jedoch typisch für das Behandlungsprotokoll. Es ist damit zu rechnen, dass viele Zwischenergebnisse ähnlichen Schwankungen unterliegen. Bei Kurvenverläufen wie in Abbildung 4.3 ist die im Protokoll verwendete Bewertung von Veränderungen beispielsweise ebenfalls recht unzuverlässig. Auch wenn eine betrachtete Variable tendenziell steigt, so können doch 15 Minuten auseinander liegende Zeitpunkte gefunden werden, deren absolute Differenz auf eine signifikante Verminderung schließen ließe.
Da die Auswirkungen häufig und eher willkürlich wechselnder diskreter Bewertungen schwer abzuschätzen sind, ist die Qualität der Vergleiche im Protokoll fragwürdig. Tiefer gehende Experimente könnten hier mehr Klarheit schaffen.

5. Diskussion

5.1. Zusammenfassung

Explizite Behandlungsprotokolle bieten der Medizin neue Möglichkeiten, die Wirksamkeit von Therapien zu validieren und Patienten die jeweils bestmögliche Versorgung zukommen zu lassen. Dem steht jedoch ein hoher Entwicklungsaufwand für die Protokolle gegenüber.
Das Thema dieser Diplomarbeit ist die exemplarische Validierung eines intensivmedizinischen Behandlungsprotokolls durch geeignete Formalisierung und Anwendung auf reale Patientendaten. Es soll dabei untersucht werden, ob durch den Einsatz geeigneter Werkzeuge und protokollierter Daten die (Weiter-)Entwicklung und Validierung solcher Protokolle in frühen Entwicklungsstadien vereinfacht wird.
Eine geeignete Operationalisierung des für diese Arbeit zur Verfügung stehenden Behandlungsprotokolls stellt den Schwerpunkt dar. Das operationale Modell vollzieht die einzelnen Schritte des Protokolls nach. Zu jeder Visite lassen sich Interventionsempfehlungen, sowie deren Ableitungen und medizinisch bedeutsame Zwischenergebnisse auf unterschiedlich abstrakten Ebenen anzeigen. Eine Datenbank-Schnittstelle ermöglicht den direkten Zugriff auf reale Patientendaten.
Nachdem das Protokoll in einem geeigneten Formalismus vorliegt, lassen sich durch geringe Anpassungen und Ergänzungen verschiedene Experimente durchführen. Da die für diese Diplomarbeit verfügbaren Datensätze neben den Patientenzuständen auch die durchgeführten Interventionen erkennen lassen, stellt die Erweiterung zu einem Kritik-System eine attraktive Möglichkeit dar, kritische Vorschläge des Modells oder kritische im Datensatz vermerkte Interventionen zu entdecken. Einem Experten auf dem Gebiet der Intensivmedizin gelang es mit Hilfe der zum Kritik-System erweiterten Wissensbasis recht zügig, einige Gründe für Unterschiede zwischen Protokollvorschlägen und Interventionen im verwendeten Datensatz ausfindig zu machen.
Ein weiteres Experiment, das ebenfalls auf das umgesetzte Modell aufsetzt abstrahierte von der Zeitpunkt-basierten Sicht des zugrundeliegenden Protokolls. Die gewonnene übersichtliche Darstellung der Vorschläge im Verlauf einer realen Behandlung half bei der Suche nach Fehlern und Schwachstellen, nicht nur beim Behandlungsprotokoll, sondern auch bei der Vorverarbeitung der Patientendaten.
Durch sukzessive Verbesserungen des Modells und eine Steigerung der Datenqualität gelang schließlich eine spürbare Verbesserung der Kritik-Ergebnisse. Bei einer abschließenden Diskussion mehrerer Fälle mit dem Experten fanden sich selbst für die von der tatsächlich durchgeführten Therapie abweichenden Vorschläge gute Gründe[1].
Bezüglich der in dieser Diplomarbeit untersuchten Methode lässt sich feststellen, dass die gemachten Experimente mit dem exemplarisch umgesetzten Protokoll ermutigende

[1] Hierbei waren Besonderheiten des Protokolls zu berücksichtigen. Abschnitt 4.2.3 skizziert die regional bedingt etwas andere Sichtweise des Protokolls.

Ergebnisse lieferten. Das Auffinden kritischer Fälle, deren Analyse und die Korrektur von Fehlern werden vom beschriebenen Modell und der Wissenserwerbsumgebung MOBAL in geeigneter Weise unterstützt.
Die Grenzen der Methode dürften bei einem weiter ausgearbeiteten und validierten Protokoll schließlich aufgrund von Unzulänglichkeiten protokollierter Daten erreicht werden. Wenn in den Datensätzen keine passenden Fälle zu einem Teilaspekt verfügbar sind, etwa weil die zu validierenden Teilaspekte im Laufe des Aufbaus der Wissensbasis komplexer werden, so muss die Validierung scheitern. Dies war im Rahmen dieser Arbeit bereits bei den Wartezeiten der Fall. Die wenigen Fälle, in denen während der Experimente Wartezeiten aufgestellt wurden, gaben wenig Aufschluss. Entweder liefen die Wartezeiten aus, ohne dass das betroffene Medikament herabgesetzt werden sollte, oder das Medikament musste zügig abgesetzt werden, was im Modell zum Abbruch von Wartezeiten führt. Die Konsequenzen eines Verstoßes gegen Wartezeiten konnten somit nicht untersucht werden. Die experimentellen Ergebnisse aus Abschnitt 4.3.2 zeigen jedoch, dass den Wartezeiten eine besondere Bedeutung zukommt.
Ein grundsätzliches Problem bei der Validierung anhand von Datensätzen liegt in mangelnder statistischer Unabhängigkeit zwischen untersuchten Eigenschaften und Auswahlkriterien. Sollen mit den in dieser Arbeit beschriebenen Methoden zwei Behandlungsalternativen A und B miteinander verglichen werden, so kann dies nur durch einen Vergleich selektierter Fälle geschehen. Fälle, bei denen nach Alternative A gehandelt wurde können mit solchen verglichen werden, bei denen gemäß Alternative B vorgegangen wurde. Dabei besteht jedoch die Gefahr, dass in den zur Verfügung stehenden Daten die resultierenden Mengen weitere spezielle Eigenschaften besitzen. Beispielsweise könnte sich in einem bestimmten klinischen Umfeld Alternative A „eingebürgert“ haben und daher in Standardfällen häufiger erwogen werden. Alternative B könnte demnach – ohne medizinische Grundlage – in höherem Grad mit kritischen Fällen korreliert sein und Alternative A unterlegen erscheinen. Wie in Abschnitt 1.2 bereits anhand von [Ohm96] skizziert wurde, gibt es bei der Validierung von Protokollen zahlreiche Störgrößen, die Ergebnisse verfälschen können. Selbst bei einer klinischen Validierung bereitet dies Probleme.
Den Ergebnissen dieser Arbeit zufolge erweisen sich bei der Entwicklung und Validierung einer Wissensbasis in frühen Stadien protokollierte Daten als geeignetes Hilfsmittel. In späteren Stadien der Entwicklung ist eine aufwendigere klinische Validierung wohl unumgänglich.

5.2. Verbesserungsvorschläge und Ausblick

Bei der Validierung des speziellen Protokolls ergaben sich einige Erkenntnisse und Ideen, welche Aspekte verbessert werden können.
Ein wiederkehrendes Prinzip des Protokolls ist, dass Abfragen und Definitionen sehr einfach und unmissverständlich aufgebaut sind. Beispielsweise werden Werte oft zeitlich punktuell erhoben und mit Konstanten verglichen. Zunächst stellt sich die Frage, ob eine zeitlich punktuelle Erhebung von Werten bei Verläufen wie in Abbildung 4.3 den Patientenzustand optimal widerspiegelt, oder ob Glättungsverfahren hilfreich sind. Wie sich bereits bei den Experimenten zeigte, erhöhen Methoden zur Artefakterkennung die Datenqualität spürbar, sind aber nicht Teil des Protokolls. Auswirkungen von Vergleichen mit absoluten Werten wurden bereits in Abschnitt 4.3.2 angesprochen. Abgesehen vom erwähnten Problem des häufigen Passierens der Vergleichsgrenze aufgrund von natürlichen Schwankungen in den Werteverläufen spiegeln die auf diese Weise definierten

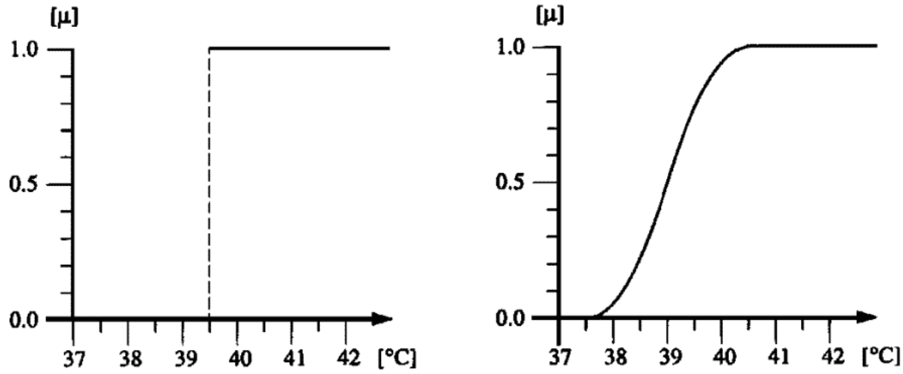

Abbildung 5.1.: Zugehörigkeitsfunktionen zum Begriff „starkes Fieber“, links als klassische, rechts als unscharfe Menge. Quelle: [Bie97]

Konzepte das menschliche Verständnis oft verzerrt wieder. Scharfe Grenzwerte wirken unnatürlich und implizieren abrupte Änderungen des Verhaltens. Durch die Verwendung unscharfer Mengen könnte eine Brücke zur menschlichen Vorstellung vieler Begriffe geschlagen werden. Als Beispiel sei der Begriff „starkes Fieber“ angeführt. Bei scharfen Grenzwerten ist die Zugehörigkeitsfunktion zwangsläufig eine Stufenfunktion wie in Abbildung 5.1 links skizziert. Trifft ein Protokoll Entscheidungen anhand derart definierter Begriffe, so sind Werte ober- und unterhalb des Grenzwertes nicht weiter unterscheidbar. Eine feinere Granularität der Ausgangsdaten wird also nicht ausgenutzt, um den Wert von 39.5 C° kann allerdings eine beliebig kleine Werteänderungen zu einer sprunghaften Änderung des Protokollverhaltens führen. Die Verwendung unscharfer Mengen erlaubt hingegen eine intuitivere Modellierung des Begriffs. Nahe beieinander liegende Werte unterscheiden sich hinsichtlich der meisten Anwendungen kaum, zunehmende Differenz sollte sich hingegen im Protokollverhalten widerspiegeln. Abbildung 5.1 rechts zeigt eine der Intuition eher entsprechende stetige Zugehörigkeitsfunktion.

Morris führt in [Mor98] gegen die Verwendung unscharfer Logik das Argument an, dass diese gegenüber der klassischen Logik zu unspezifisch sei. Dem kann entgegen gehalten werden, dass sie ebenso formal und exakt wie die klassische Logik ist, lediglich komplexere Begriffsdefinitionen zulässt. Zu erzwingen, dass Begriffe jederzeit ganz oder gar nicht zutreffen, bringt für die Standardisierung medizinischer Begriffe Vorteile. Um flexibel auf die Änderungen von Herz-/Kreislaufparametern reagieren zu können und um von Experten verwendete Begriffe und die daran geknüpften Handlungen in Form eines operationalen Protokolls umzusetzen ist unscharfe Logik hingegen vielversprechender.

Zu den im Protokoll auffällig einfach umgesetzten Aspekten gehört weiterhin die Art und Weise, wie einige Variablen aus anderen streng hierarchisch abgeleitet werden. Der hierarchiehöchste Wert ist dabei der einzig relevante. Ist dieser verfälscht, eventuell aufgrund von Messfehlern, so wird der abgeleitete Wert gleichermaßen verfälscht. Aktuellere Werte mit niedrigerer Priorität werden ignoriert, solange ein älterer Wert höherer Priorität noch gültig ist.

Die Gültigkeit von Werten ist ebenso einfach definiert. Werte sind bis zu einem bestimmten „Alter“ uneingeschränkt gültig, danach werden sie ungültig.

Zur Verbesserung bieten sich Methoden an, die robust gegen verrauschte Daten sind, die Redundanz mehrerer Werte ausnutzen und die schwindende Angemessenheit älterer

Werte zur Beurteilung des Patientenzustands berücksichtigen. Ein erster Schritt wäre eine nach Zuverlässigkeit, Alter und Relevanz gewichtete Komposition der einzelnen Variablen. Dies würde besonders bei veraltenden Werten zu einem stetigerem Verhalten führen, als die Vergleiche im Protokoll.
Bezüglich der Stetigkeit der Protokollvorschläge fällt weiterhin auf, dass diese bei jeder Visite von Anfang an neu berechnet werden. Die Ergebnisse aus Abschnitt 4.3.2 zeigen, dass bei der blutdruckgesteuerten Gabe von Vasokonstringenzien auf Wartezeiten nur schlecht verzichtet werden kann, da andernfalls die Gefahr einer oszillierenden Änderung der Medikation besteht. Wartezeiten stellen ein strategisches Element dar, mit dem für das Zeitpunkt-orientierte Konzept längerfristig widersprüchliche Therapieziele ausgeschlossen werden.
Darüber hinaus erscheint es grundsätzlich sinnvoll, auf einer abstrakteren Ebene einen Behandlungsplan vorzusehen und Therapieziele längerfristig zu verfolgen. Während im Protokoll in jeder Visite der aktuell durchzuführende Schritt fast unabhängig von vorangegangenen Entscheidungen abgeleitet wird, können bei einem Planungsszenario Ziele so lange aufrecht erhalten werden, bis sie erreicht sind, oder bis ein explizit formuliertes Abbruchkriterium erfüllt ist, was zur Erstellung eines neuen Therapieplans führt. Bei Notfällen kann der bestehende Plan eventuell nur kurzfristig außer Kraft gesetzt werden, bis unter normalen Bedingungen fortgesetzt werden kann.
Im Kontext eines bestehenden Therapieziels können die Berechnungen sich unmittelbar auf die Wahl der konkreten Mittel zu dessen Erreichung beschränken. Die Fortführung einmal aufgestellter Behandlungspläne ist auf diese Weise auch dann möglich, wenn die ursprüngliche Ursache für die Wahl einer Therapie kurzfristig nicht mehr zu beobachten ist. Dies hilft kurze sprunghafte Änderungen im Behandlungsablauf zu verhindern und somit die Stetigkeit zu erhöhen.
Die Eigenschaft abstrakter Therapien des Modells, bestimmte Vorbedingungen und Auswirkungen zu besitzen, stellt eine günstige Grundlage für Planung dar. Klingspor stellt in [Kli98] die Planung anhand operationaler Begriffe vor. Diese integrieren Vorbedingung und Handlung und bilden die Operatoren der Planungsebene. Im intensivmedizinischen Kontext könnten die abstrakten Therapien diese Operatoren bilden. Die Nachbedingungen könnten sich an zuvor explizierten Auswirkungen der Therapien auf die verschiedenen Parameter orientieren. Auf diese Weise würde das Modell in stärkerem Maße auf repräsentierten kausalen Zusammenhängen statt auf heuristischen Assoziationen beruhen. Die Entwicklung eines Protokolls für die bestmögliche Behandlung könnte auf diese Weise von einer zunehmend genauen Modellierung des medizinischen Sachbereichs profitieren.
Schließlich fällt auf, dass die Beschränkung auf Visiten dem klinischen Alltag aus Medizinersicht nachempfunden ist. Bei der rechnerunterstützen Behandlung ist hingegen eher eine Ausdehnung der derzeit bereits üblichen kontinuierlicher Überwachung relevanter Parameter („Monitoring“) erstrebenswert. Die Ausgabe von Interventionsvorschlägen sollte kontinuierlich als Reaktion auf bestimmte Ereignisse möglich sein, statt sich an pauschalen Zeitspannen zwischen Visiten zu orientieren.
Das untersuchte Behandlungsprotokoll bietet noch auf vielen Gebieten Raum für Veränderungen und Experimente. Eine stärkere Anpassung an die Gegebenheiten deutscher Kliniken, die systematische Untersuchung verschiedener Vorverarbeitungsoperatoren und die Modellierung mit anderen als den im Protokoll angewandten Methoden stehen noch aus und können Gegenstand nachfolgender Arbeiten sein.

Anhang A – Technische Details

Dieser Abschnitt beschreibt einige technische Aspekte der Umsetzung des Modells, die im Kontext der Diplomarbeit von untergeordneter Bedeutung sind. Dabei handelt es sich einerseits um die Erweiterung der Funktionalität von MOBAL durch externe Prädikate und eine Datenbank-Anbindung, andererseits um Maßnahmen, die halfen, die Performanz des System deutlich zu steigern.

Verwendung externer Prädikate in MOBAL

Da der MOBAL-Formalismus allein für einige Berechnungen und Anfragen an die Datenbank nicht angemessen ist, war eine externe Schnittstelle erforderlich. Eine in [Kie] dokumentierte Schnittstelle stellt das „MOBAL Program Interface" dar. Diese erlaubt „von außen" auf die Ableitung gewisser Fakten zu reagieren und beispielsweise andere Fakten als Antwort in die Faktenbasis einzufügen. Direkter erwies sich jedoch die Verwendung externer Prädikate, die auch als Schnittstelle von PROLOG zu anderen Programmiersprachen üblich sind. Dabei werden Prämissen bestimmter Prädikate generell extern ausgewertet und die Ergebnisse einfach durch Variablen-Substitutionen bzw. das Fehlschlagen der Anfrage übergeben.
Um diese Funktionalität in MOBAL einzubinden bedarf es einer Erweiterung der Inferenzmaschine, bei der, ähnlich wie bei den autoepistemischen Operatoren, bestimmte Prädikate eine besondere Verarbeitung anstoßen. Die „IM2-DB" stellt den Operator extern(...) zur Verfügung. Dieser bekommt eine positive MOBAL-Prämisse als Argument übergeben, die nicht wie gewöhnlich ausgewertet wird, sondern über das gleichnamige PROLOG-Prädikat des Moduls mextern.
Nachdem MOBAL mit der IM2-DB kompiliert war, ließen sich die externen Prädikate einfach in PROLOG realisieren. Eine Datei mextern.pl definiert das Modul und lässt sich über den „Prolog-Listener" von MOBAL mit Hilfe des Kommandos

```
use_module('mextern.pl').
```

einbinden.

Datenbankanbindung an MOBAL

Zu den wichtigsten Anwendungen der externen Schnittstelle gehört die Datenbank-Anbindung. Diese gelang mit Hilfe des Oracle Moduls von RDTDB, das als „externes Tool" aus dem Verzeichnis rdtdb/4.0 nachgeladen werden muss. Die Benutzerdaten für die Datenbank können über das Oracle-Menü (Menüpunkt „Parameter-Display") eingegeben werden. In der aktuellen Umsetzung werden sie automatisch bei jeder Initialisierung des Moduls mextern eingetragen.
Der Quellcode des Oracle-Moduls ist ausführlich dokumentiert. Es enthält ein Prädikat oracle:oracle_process_sql_term/3, das SQL-Anfragen verarbeitet und die Ergebnisse als

Listen zurück liefert. Entsprechend werden Anfragen über einen SQL-Generator aufgebaut und die Ergebnisse im Kontext der jeweiligen Anfrage zurückgeliefert. Zu beachten war, dass die Reihenfolge der Ergebnisse aufgrund der Verarbeitung in PROLOG invertiert ist, und dass numerische Rückgabewerte auf ganze Zahlen gerundet werden, auch wenn zwei Nachkommastellen angezeigt werden. Bei SQL-Anfragen, die die Gleichheit von Dezimalbrüchen testen, stellt sich die Problematik einer ungenauen Darstellung als binäre Gleitkommazahl. Ob ein Attributwert gleich 0.07 ist, ist beispielsweise nicht unmittelbar zu erfragen, da dieser Wert in der internen Repräsentation nicht existiert und daher auf einen minimal größeren Wert abgebildet wird. Um dieses Problem zu umgehen werden die letzten Nachkommastellen zum Runden (in ASCII-Darstellung) verwendet.

Caching in MOBAL

Aufgrund des Volumens der verarbeiteten Daten waren Strategien zur Steigerung der Performanz von besonderem Interesse. Es zeigte sich bei Testläufen, dass besonders die hohe Zahl von Datenbank-Anfragen das System ausbremste, wobei sich bestimmte Anfragen in leichter Variation ständig wiederholten. Beispielsweise wurden für noch nicht verabreichte Medikamente wiederholt die letzten Änderungen oder oszillierenden Verläufe gesucht. Besonders massive Performanzprobleme stellten sich bei einer höheren Rückwärtsinferenz-Tiefe ein, da in diesem Fall sogar oft mehrfach hintereinander die gleiche Anfrage gestellt wurde.

Die erste Maßnahme war, die letzten SQL-Anfragen und deren Ergebnisse im Speicher zu halten um sie bei wiederholten Anfragen schneller beantworten zu können. Als besonders effektive Maßnahme erwies sich auch die Ableitung der inferentiellen Hülle ausschließlich über Vorwärtsinferenz. Darüber hinaus wurde die Möglichkeit eingeräumt häufig abgefragte Attribute einzelner Patienten, die sich selten ändern, in einer kompakten Intervallform im Hauptspeicher zu cachen.

Nur zum Vergleich der Größenordnung seien die folgende Rechenzeiten genannt: Während bei frühen Versuchen mit ungebundenen Zeitpunktvariablen und Rückwärtsinferenz Rechenzeiten von über 30 Minuten pro Visite vorkamen, lagen sie bei geschickterer Repräsentation, ohne Rückwärtsinferenz und ohne Caching bei mehreren Minuten, und schließlich mit Caching unter 30 Sekunden.

Literaturverzeichnis

[AE97] ABDOLVAHAB-EMMINGER, HAMID: *Physikum exakt.* Georg Thieme Verlag Stuttgart/New York, 1997.

[BGGP93] BAMBERGER, S., U. GAPPA, K. GOOS und K. POECK: *Teilautomatische Wissenstransformation zur Unterstützung der Wissensakquisition.* In: *Proceedings of XPS*, 1993.

[BGI99] BAUER, M., U. GATHER und M. IMHOFF: *The Identification of Multiple Outliers in Online Monitoring data.* Technischer Bericht29, SFB 475, 1999.

[Bie97] BIEWER, B.: *Fuzzy-Methoden.* Springer Verlag, 1997.

[Bis95] BISKUP, J.: *Grundlagen von Informationssystemen.* Vieweg Verlag, 1995.

[Bra87] BRATKO, I.: *PROLOG: Programmierung für künstliche Intelligenz.* Addison-Wesley Verlag, 1987.

[CCPR00] CALVELO, D., M.-C. CHAMBRIN, D. POMORSKI und P. RAVAUX: *Towards symbolization using data-driven extraction of local trends for ICU monitoring.* Artificial Intelligence in Medicine, 19, 2000.

[Cla85] CLANCEY, W.: *Heuristic Classification.* Artificial Intelligence, 27, 1985.

[Cla86] CLANCEY, W.: *From GUIDON to NEOMYCIN and HERACLES in Twenty Short Lessons.* AI Magazine, 7(3):289 – 350, 1986.

[ES] E. SOMMER, W. EMDE, J.-U. KIETZ S. WROBEL: *Mobal 4.1b9 User Guide – Draft.* GMD – German National Reasearch Center for Information Technology, St. Augustin.

[Hil98] HILDEBRANDT, H. (Herausgeber): *Pschyrembel, Klinisches Wörterbuch.* Walter de Gruyter, 258. Auflage, 1998.

[IBGL97] IMHOFF, M., M. BAUER, U. GATHER und D. LÖHLEIN: *Statistical Pattern Detection in Univariate Time Series of Intensive Care On-Line Monitoring Data.* Intensive Care Med, 24, 1997.

[Imh00] IMHOFF, M.: *Medical Errors and How to Prevent Them.* http://www.frice.nl/Editorials/Medical_Errors_MI_Dec_2000.htm, 2000.

[KCD00] KOHN, L. T., J. M. CORRIGAN und M. S. DONALDSON (Herausgeber): *To Err Is Human: Building a Safer Health System.* National Acadamy Press, Washington, 2000.

[Kie] KIETZ, J.-U.: *Mobal's Program Interface.* GMD.

[Kli98] KLINGSPOR, V.: *Reaktives Planen mit gelernten Begriffen.* Doktorarbeit, Universität Dortmund, Fachbereich Informatik, 1998.

[KMSK92] KOSKI, E., A. MÄKIVIRTA, T. SUKUVAARA und A. KARI: *Development of an expert system for haemodynamic monitoring: computerized symbolization of on-line monitoring data.* International Journal of Clinical Monitoring and Computing, 8:289 – 293, 1992.

[Ma95] MA, HONGYING: *Mapping clause of Arden Syntax with HL7 and ASTM E 1238-88 standard.* International Journal of Bio-Medical Computing, 38:9 – 21, 1995.

[MBJ99] MORIK, K., P. BROCKHAUSEN und T. JOACHIMS: *Combining statistical learning with a knowledge–based approach – A case study in intensive care monitoring.* In: *International Conference on Machine Learning*, 1999.

[Müh94] MÜHLENBROCK, M.: *Menschliches und Maschinelles Lernen – Bildung und Revision von Begriffsstrukturen zum Tag/Nachtzyklus.* Diplomarbeit, Universität Dortmund, Lehrstuhl Informatik VIII, 1994.

[MIB+00] MORIK, K., M. IMHOFF, P. BROCKHAUSEN, T. JOACHIMS und U. GATHER: *Knowledge discovery and knowledge validation in intensive care.* Artificial Intelligence in Medicine, 19, 2000.

[Mik94] MIKSCH, S.: *Wissensbasierte Systeme in der Intensivmedizin: Was können sie, was sollten sie können?* In: *EDV im Krankenhaus.* Springer Verlag Wien, 1994.

[Mil56] MILLER, G.: *The magical number seven plus or minus two – some limits on our capacity for processing information.* Psychological Review, 63:81 – 97, 1956.

[Mor89] MORIK, K.: *Sloppy Modelling.* In: MORIK, K. (Herausgeber): *Knowledge Representation and Organization in Machine Learning*, Band 347 der Reihe *Lecture Notes in Artificial Intelligence*, Seiten 107 – 134. Springer Verlag, 1989.

[Mor98] MORRIS, A.: *Principles and Practice of Intensive Care Monitoring*, Kapitel 77, Seiten 1355 – 1381. McGraw–Hill, 1998.

[Mut97] MUTSCHLER, E.: *Arzneimittelwirkungen, Lehrbuch der Pharmakologie und Toxikologie.* Wissenschaftliche Verlagsgesellschaft mbH Stuttgart, 7. Auflage, 1997.

[MWKE93] MORIK, K., S. WROBEL, J.-U. KIETZ und W. EMDE: *Knowledge Acquisition and Machine Learning: Theory, Methods, and Applications.* Academic Press, 1993.

[Net69] NETTER, F. H.: *The CIBA collection of medical illustrations*, Band 5. 1969.

[New82] NEWELL, A.: *The Knowledge Level.* Artificial Intelligence, (18):87 – 127, 1982.

[NS97] NERODE, A. und R. A. SHORE: *Logic for Applications.* Springer Verlag, 2. Auflage, 1997.

[Ohm96] OHMANN, C.: *MEDWIS Arbeitskreis Evaluation – Leitfaden zur Validierung von Wissensbasen.* http://www.uni-duesseldorf.de/WWW/MedFak/TheoChir/ak_eval/lf_main.htm, 1996.

[OKKN00] OWSNICKI-KLEWE, B., K.V.LUCK und B. NEBEL: *Wissensrepräsentation und Logik – Eine Einführung.* In: GÖRZ, ROLLINGER, SCHNEEBERGER (Herausgeber): *Handbuch der Künstlichen Intelligenz.* Oldenbourg Verlag, 3. Auflage, 2000.

[PGPB96] PUPPE, F., U. GAPPA, K. POECK und S. BAMBERGER: *Wissensbasierte Diagnose- und Informationssysteme.* Springer Verlag, 1996.

[PSS00] PUPPE, F., H. STOYAN und R. STUDER: *Knowledge Engineering.* In: GÖRZ, ROLLINGER, SCHNEEBERGER (Herausgeber): *Handbuch der Künstlichen Intelligenz.* Oldenbourg Verlag, 3. Auflage, 2000.

[Pup90] PUPPE, F.: *Problemlösungsmethoden in Expertensystemen.* Studienreihe Informatik. Springer Verlag, 1990.

[Pup99] PUPPE, F.: *Meta Knowledge for Extending Diagnostic Consultation to Critiquing Systems.* In: *Proceedings of EKAW-99 11th European Workshop on Knowledge Acquisition, Modeling and Management*, Seiten 367 – 372. Springer Verlag, 1999.

[SLDE97] SPENCER, R. G., C. S. LESSARD, F. DAVILA und B. ETTER: *Self-Organising Discovery, Recognition, and Prediction In The Intensive Care Unit.* Medical & biological engineering & computing, 35(2):117, 1997.

[SM96] SHAHAR, Y. und M. MUSEN: *Knowledge-Based Temporal Abstraction in Clinical Domains.* Artificial Intelligence in Medicine, 8(3):267 – 298, 1996.

[Str00] STRUSS, P.: *Modellbasierte Systeme und qualitative Modellierung.* In: GÖRZ, ROLLINGER, SCHNEEBERGER (Herausgeber): *Handbuch der Künstlichen Intelligenz.* Oldenbourg Verlag, 3. Auflage, 2000.

[The99] *Therapierichtlinien.* Chirurgische Intensivstation, Städtische Kliniken Dortmund, 1999.

[TKM00] TSIEN, C. L., I. S. KOHANE und N. MCINTOSH: *Multiple signal integration by decision tree induction to detect artifacts in the neonatal intensive care unit.* Artificial Intelligence in Medicine, 19, 2000.

[TMV91] THEWS, G., E. MUTSCHLER und P. VAUPEL: *Anatomie, Physiologie, Pathophysiologie des Menschen.* Wissenschaftliche Verlagsgesellschaft mbH Stuttgart, 4. Auflage, 1991.

[Wai98] *WAIN - Wissensbasis für ein allgemeines Arzneimittelinformationssystem mit Schwerpunkt Intensivmedizin und Nephrologie.* http://medweb.uni-muenster.de/institute/imib/forschung/projekteinf/bauchschmerz/wain2.html, 1998.

*„It may be part of human nature to err,
but it is also part of human nature to
create solutions, find better alternatives
and meet the challenges ahead.“ ([KCD00])*

Erklärung

Hiermit erkläre ich, Martin Scholz, die vorliegende Diplomarbeit mit dem Titel

Modellierung eines intensivmedizinischen Behandlungsprotokolls zur Validierung anhand realer Patientendaten

selbständig verfasst und keine anderen als die hier angegebenen Hilfsmittel verwendet zu haben.

Dortmund, 18. Juli 2001