

WIE KANN DER PROZESS DER COMPUTERISIERUNG VON MEDIZINISCHEN LEITLINIEN UNTERSTÜTZT WERDEN?

Kaiser K^{1,2}, Miksch S^{1,3}

Kurzfassung

Die Formalisierung von medizinischen Leitlinien in ein computer-interpretierbares Format ist eine große Herausforderung an MedizinerInnen und InformatikerInnen. In dieser Arbeit wollen wir zeigen, wie wir diesen Prozess durch Anwenden von NLP-Methoden automatisieren können. Mit den entwickelten Methoden können die notwendigen Informationen identifiziert, Abläufe definiert und visualisiert werden. Computer-interpretierbare Leitlinien können in klinische Informationssysteme integriert werden und das klinische Personal bei der Behandlung von PatientInnen unterstützen.

Abstract

Formalising clinical practice guidelines in a computer-interpretable format is a great challenge for both medical and computer-science experts. In this work we show how this process can be automated using NLP techniques. With the developed methods we can identify the necessary information, define and visualise processes. Computer-interpretable clinical practice guidelines can be integrated in clinical information systems and support the health care personnel in patient care.

Keywords – computerized clinical practice guidelines, NLP, UMLS, semantic relations

1. Einleitung

Medizinische Leitlinien sind wichtige Instrumente, die evidenzbasiertes medizinisches Wissen beinhalten und als praxisorientierte Entscheidungshilfen zur Verfügung stellen. Durch die Integration von medizinischen Leitlinien in den klinischen Arbeitsprozess und besonders in PatientInnendaten-Managementssysteme, werden insbesondere die Bedürfnisse der PatientInnen unterstützt und die medizinische Leistung auf einer individuellen Basis erhöht. Um dies erreichen zu können, wurden s.g. Leitlinien-Repräsentationssprachen entwickelt (siehe [6] für einen Überblick). Die Erstellung eines formalen Modells einer Leitlinie in einer solchen Sprache stellt eine große Herausforderung dar: sowohl medizinisches Wissen, als auch Wissen über die formalen Anforderungen der Sprache sind notwendig. Weiters hat die bisherige Erfahrung gezeigt, dass eine Formalisierung in einem Schritt sehr komplex und schwierig bis gar unmöglich ist. Aus diesen Gründen wurden s.g. Abstraktionssprachen entwickelt, die als Basis für eine weitere Formalisierung in ein endgültiges

¹ Institut für Softwaretechnik und Interaktive Systeme, Technische Universität Wien

² Zentrum für Medizinische Statistik, Informatik und Intelligente Systeme, Medizinische Universität Wien

³ Department für Information und Knowledge Engineering, Donau-Universität Krems

computer-ausführbares Leitlinienformat dienen (z.B. MHB, die „Many-Headed Bridge“ zwischen Leitlinien-Formaten [9] oder GEM, das „Guideline Elements Model“ [10]). Trotzdem werden sowohl der Abstraktions- als auch der endgültige Modellierungsprozess noch gänzlich manuell durchgeführt [8]. Der Trainingsaufwand ist in beiden Fällen enorm und der eigentliche Modellierungsprozess zeit- und kostenintensiv.

Aus diesem Grund stellen wir eine Methodik vor, die zumindest Teile des Abstraktionsprozesses automatisiert, indem beispielsweise Ausschnitte aus der Leitlinie automatisch identifiziert werden und in weiterer Folge in ein abstraktes Leitlinienformat transformiert werden können. In unserem Fall verwenden wir *MHB* [9], als Abstraktionsformat. Zur Durchführung der Automatisierung greifen wir auf NLP-Methoden zurück und verwenden das UMLS¹), um relevante medizinische Konzepte und ihre Semantik identifizieren zu können. Somit können wir relevante Information extrahieren und im MHB-Format darstellen. Ein weiterer bedeutender Teil des Abstraktionsprozesses ist, den Ablauf der einzelnen Aktionen zu bestimmen, die in der Leitlinie beschrieben werden. Somit gehen wir einen Schritt weiter als bisherige Ansätze, die linguistische Methoden nur als Preprocessing des eigentlichen Modellierungsprozesses angewandt haben (z.B. [7]).

In der vorliegenden Arbeit stellen wir unsere Gesamtidee dar, sowie die Kernkomponenten des Systems, das wir entwickeln. Außerdem diskutieren wir die Vorteile einer solchen Methodik und wie sie zur Verbreitung von computerisierten medizinischen Leitlinien beitragen kann.

2. Methodik

Die Umsetzung von Text-Leitlinien in eine formale Darstellung ist eine komplexe Aufgabe. Verschiedene AkteurInnen sind dabei in den Prozess involviert. Obwohl verschiedene Programme und Werkzeuge entwickelt wurden, um diesen Prozess zu unterstützen (siehe [1] für einen Überblick), ist es eine sehr schwierige und komplexe Aktivität, die noch immer gänzlich manuell durchgeführt wird. Da wir Teile des Formalisierungsprozesses automatisieren wollen, haben wir in einem ersten Schritt die Anforderungen an ein solches System analysiert und können die Ergebnisse folgendermaßen zusammenfassen:

- Die Informationsdimensionen in einer Leitlinie müssen entsprechend den Vorgaben von MHB automatisch identifiziert werden.
- Die Abfolge der einzelnen Therapieschritte soll automatisch bestimmt werden.
- Das MHB-Modell muss erstellt werden. Dabei wollen wir einen dokumenten-zentrierten Ansatz anwenden, bei dem die formalen Modelle der Informationsdimensionen mit den entsprechenden Textstellen im Original-Dokument verbunden werden.
- Die Formalisierung muss nachvollziehbar und verständlich sein. Dies wollen wir durch entsprechende Visualisierungsmethoden unterstützen.
- Die Formalisierung muss editierbar sein, um Informationen hinzufügen oder abändern zu können.
- Das System muss ein Versionsmanagement enthalten. Leitlinien werden ständig aktualisiert. Diese Aktualisierungen müssen auch in den computerisierten Formaten widergespiegelt werden.

Basierend auf dieser Analyse haben wir unser System entwickelt (*Abbildung 1* stellt die Architektur dar).

¹ Unified Medical Language System, <http://umlsinfo.nlm.nih.gov/> (Letzter Zugriff: 1. Dez. 2009)

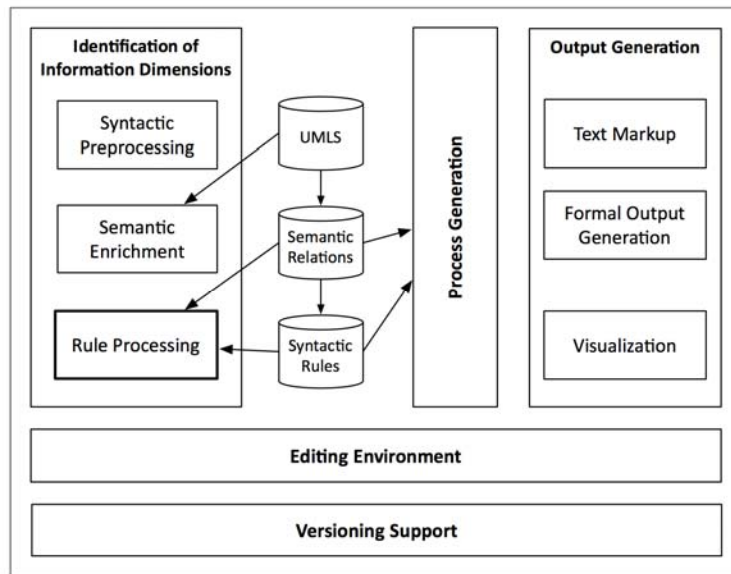


Abbildung 1: Architektur des Leitlinien-Formalisierungssystems.

Das System besteht aus vier Hauptmodulen. Es verwendet Regelbasen und das UMLS als medizinisches Terminologiesystem. Wir werden nun die Module beschreiben und dabei speziell auf die Aufgaben, die diese entsprechend den zuvor definierten Anforderungen zu erfüllen haben, eingehen.

2. 1. Identifizierung von Informationsdimensionen

Die Aufgabe dieses Modules ist es, die Informationsdimensionen im Text des Leitliniendokumentes entsprechend den Vorgaben von MHB zu identifizieren. Wir haben dabei semantische Relationen aus dem UMLS Semantic Network identifiziert, die bestimmte Informationsdimensionen beschreiben können. Beispielsweise können wir im Satz „*Women may eat a light diet in established labor*“ die Relation ‚*Population Group performs Therapeutic or Preventive Procedure*‘ erkennen, die eine Aktion beschreibt. Diese Relationen werden in syntaktischen Regeln abgebildet, um sie mittels NLP aus dem Text extrahieren zu können. Die Regeln sind in der *tgrep*-Syntax geschrieben und werden auf syntaktischen und Part-Of-Speech (POS)-annotierten Text angewandt. Da die Ergebnisse des Parsing-Prozesses, der syntaktische und POS-Informationen generiert, manchmal fehlerhaft sind, verwenden wir noch ein *Syntactic Preprocessing* Modul.

2. 2. Bestimmung des Prozessablaufes

Das vorhergehende Modul hat die unterschiedlichen Informationsdimensionen in einem Leitlinien-Dokument bestimmt. Unter anderem hat es auch Aktionen erkannt, die vom Gesundheitspersonal auszuführen sind, sowie Bedingungen, die die Ausführung der Aktionen kontrollieren. Der Großteil der Aktionen steht allerdings noch immer in keiner Verbindung zu einer anderen Aktion. Das Finden der Verbindungen zwischen den Aktionen ist die Aufgabe dieses Modules. Dabei ist es nicht nur wichtig die Reihenfolge des Ablaufes von Prozessen zu erkennen, sondern auch, ob es sich um Teilprozesse oder alternative Aktionen handelt. Um Teilprozesse oder alternative Aktionen zu identifizieren verwenden wir semantische Informationen, die uns das UMLS liefert. Schwieriger ist es, die einzelnen Aktionen zu ordnen. Durch die erzählerische Form der Leitlinie gehen wir grundsätzlich von einem sequentiellen Ablauf der einzelnen Aktionen aus. Um Alternativen zu erkennen, benutzen wir aber wiederum semantische Information aus dem UMLS.

2.3. Erstellen der Ausgabeformate

Die Information aus den vorigen Modulen muss nun in das MHB-Format übertragen werden. Die Syntax von MHB basiert auf XML. Das Format wird noch immer von Nicht-InformatikerInnen verstanden, ist aber trotzdem weitaus formaler als freier Text.

Sowohl die XML-Darstellung als auch das originale Leitlinien-Dokument sind nicht geeignet, um einen Überblick über die gesamte Therapie, die die Leitlinie beschreibt, zu geben. Deshalb ist es notwendig die Leitlinie mit passenden Mechanismen zu veranschaulichen. Es hat sich gezeigt, dass Informationsvisualisierungstechniken in der Lage sind, die Logik eines Therapieplanes adäquat zu kommunizieren [2]. Beispielsweise können flowchart-ähnliche Visualisierungen (z.B. wie in Protégé [3]) oder visuelle Metaphern (z.B. in [5]) verwendet werden (siehe *Abbildung 2*). Zusätzlich können Interaktions- sowie Fokus+Kontext-Mechanismen die Anwender unterstützen. Wir planen in unser Framework *CareVis*-Komponenten (siehe [1]) zur Visualisierung zu integrieren.

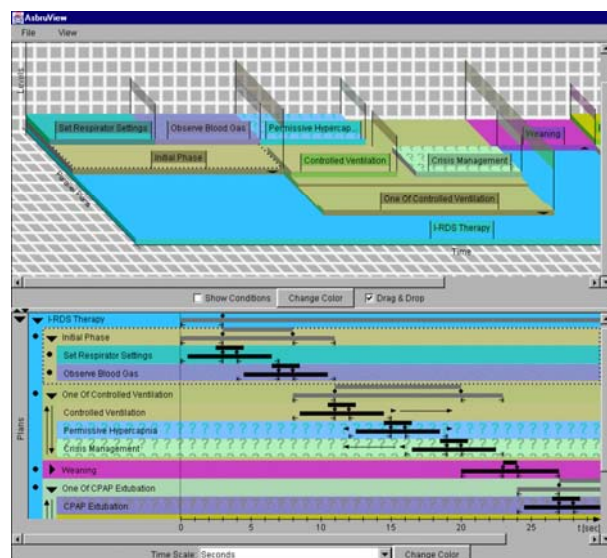


Abbildung 2: Beispiele für eine metaphor- bzw. PlanningLine-basierte Visualisierungen.

2.4. Editier-Umgebung

Die automatisch erstellten Ausgaben der vorhergehenden Module müssen überprüft und – wenn notwendig – korrigiert werden. Außerdem müssen Inkonsistenzen behoben werden und fehlende Information im Modell muss erkannt und hinzugefügt werden. Fehlende Information ist häufig dort beobachtbar, wo bestimmte Aspekte einer Domäne als zu trivial befunden werden, um sie für MedizinerInnen zu wiederholen. Allerdings sind diese Informationen für ein formales und automatisch ausführbares Modell notwendig und müssen eingefügt werden. Durch Abklärung dieser Sachverhalte und durch den Erwerb der fehlenden Information zu diesem frühen Zeitpunkt kann wertvolle Zeit im weiteren Modellierungsprozess gespart werden.

2.5. Versionsmanagement

Leitlinien haben nur eine begrenzte Gültigkeit. Da sich das medizinische Wissen verändert, müssen auch die medizinischen Leitlinien an den Stand der Technik angepasst werden. Diese Aktualisierungen müssen sich auch in der computerisierten Form der Leitlinien widerspiegeln. In [4]0 haben

wir gezeigt, wie Änderungen in unterschiedlichen Versionen automatisch erkannt werden und wie die entsprechenden formalen Modelle abgeändert werden müssen.

3. Bisherige Ergebnisse

Ein Großteil der bisherigen Arbeit war das Erstellen der Regelbasis. Wir haben drei Leitlinien aus den Bereichen Onkologie, Geburtshilfe und Diätetik verwendet, um Regeln zu erlernen. Bisher konnten wir mit unserer Methode eine Regelbasis für die Identifizierung der MHB-Informationsdimensionen entwickeln, die bereits zu vielversprechenden Ergebnissen geführt hat. Um den Prozessablauf zu bestimmen, werden ebenfalls syntaktische und semantische Regeln verwendet. Wir können bisher Subprozesse und alternative Prozesse identifizieren. Bei letzteren ist es allerdings aufgrund von unklaren Formulierungen im Text oft schwierig eine eindeutige Lösung zu finden. Dies kann nur durch Rücksprache mit medizinischen ExpertInnen gelöst werden.

Das Erstellen der formalen Leitlinie und das Markieren der Textbausteine im Originaltext sind bereits realisiert. Demnächst wird mit der Implementierung von Visualisierungsmethoden basierend auf MHB begonnen. Dabei werden wir Methoden wie in *CareVis* (siehe [1]) entsprechend adaptieren. Diese sollen in die Editier-Umgebung integriert werden, um eine manuelle Korrektur einfach und intuitiv durchführen zu können. Für das Versionsmanagement werden wir die in [4] vorgestellte Methode an MHB angepasst.

4. Zusammenfassung und Diskussion

Mit der vorgestellten Methodik und dem System, das diese implementiert, gehen wir einen wichtigen Schritt in Richtung computerisierte Leitlinie. Durch den Abstraktionsschritt können wir Unklarheiten, fehlende oder implizit formulierte Information, verschiedene Interpretationsmöglichkeiten oder Abweichungen in Dokumenten, die dasselbe vermitteln sollen, schon in einem frühen Stadium erkennen und beheben. Weiters hat sich gezeigt, dass das Erstellen eines computerisierbaren Leitlinienmodells wesentlich einfacher aus einem MHB-Modell möglich ist, als aus der Text-Leitlinie direkt [9].

Mit der Automatisierung von Teilen des Abstraktionsprozesses entlasten wir die AkteurInnen, die in den Prozess involviert sind: medizinische ExpertInnen und InformatikerInnen. Trotzdem bleibt der Prozess nachvollziehbar und verständlich: einerseits durch das Verbinden der Textbausteine in der Originalleitlinie mit den jeweiligen MHB-Modellen, andererseits durch die interaktiven Visualisierungsmethoden, die es erlauben sowohl einen Überblick als auch Details des gesamten Leitlinien-Prozesses zu kommunizieren. Durch die Editierbarkeit der erstellten Informationen können nicht nur Fehler in den automatisch erstellten Modellen korrigiert werden, sondern auch fehlende Information hinzugefügt werden.

Weiters hoffen wir, dass wir den weiteren Formalisierungsschritt (d.h. die Transformation von einem MHB-Modell in eine computer-interpretierbare Leitliniensprache, wie Asbru oder PROforma) ebenfalls automatisieren können. Dadurch können Leitlinien schneller und effizienter in den klinischen Arbeitsablauf (z.B. mit einem PatientInnendaten-Managementsystem) integriert werden. Das medizinische Personal wird bei der Behandlung der PatientInnen optimal unterstützt, da die empfohlenen Behandlungsvorschläge individuell auf die PatientInnen abgestimmt sind.

5. Danksagung

Dieses Projekt wurde teilweise vom „Fonds zur Förderung der wissenschaftlichen Forschung FWF“ (Austrian Science Fund), Projektnummer L290-N04, unterstützt.

6. Literatur

- [1] AIGNER, W., KAISER, K., MIKSCH, S., Visualization techniques to support authoring, execution, and maintenance of clinical guidelines. In *Computer-based Medical Guidelines and Protocols: A Primer and Current Trends*, volume 139 of *Studies in Health Technology and Informatics*, chapter 8, pages 140–159. IOS Press, 2008.
- [2] AIGNER, W., MIKSCH, S., Communicating the logic of a treatment plan formulated in Asbru to domain experts. In *Proceedings of the Symposium on Computerized Guidelines and Protocols (CGP 2004)*, volume 101 of *Studies in Health Technology and Informatics*, pages 1–15, Prague, Czech Republic, 2004. IOS Press.
- [3] GENNARI, J.H., MUSEN, M.A., FERGERSON, R.W., GROSSO, W.E., CRUBÉZY, M., ERIKSSON, H., NOY, N.F., TU, S.W., The Evolution of Protégé: An Environment for Knowledge-based Systems Development. *International Journal of Human Computer Studies*, 58(1):89–123, 2003.
- [4] KAISER, K., MIKSCH, S., Versioning computer-interpretable guidelines: Semi-automatic modeling of 'living guidelines' using an information extraction method. *Artificial Intelligence in Medicine*, 46(1):55–66, 2009.
- [5] KOSARA, R., MIKSCH, S., Metaphors of Movement: A Visualization and User Interface for Time-Oriented, Skeletal Plans. *Artificial Intelligence in Medicine*, Special Issue: Information Visualization in Medicine, 22(2):111–131, May 2001.
- [6] PELEG, M., TU, S.W., BURY, J., CICCARESE, P., FOX, J., GREENES, R.A., HALL, R. JOHNSON, P.D., JONES, N., KUMAR, A., MIKSCH, S., QUAGLINI, S., SEYFANG, A., SHORTLIFFE, E.H., STEFANELLI, M., Comparing computer-interpretable guideline models: A case-study approach. *Journal of the American Medical Informatics Association (JAMIA)*, 10(1):52–68, Jan-Feb 2003.
- [7] SERBAN, R., TEN TEIJE, A., VAN HARMELEN, F., MARCOS, M., POLO-CONDE, C., Extraction and use of linguistic patterns for modelling medical guidelines. *Artificial Intelligence in Medicine*, 39(2):137–149, 2007.
- [8] SEYFANG, A., KAISER, K., MIKSCH, S., Modelling clinical guidelines and protocols for the prevention of risks against patient safety. In *Proceedings of the XXII International Conference of the European Federation for Medical Informatics*, volume 150 of *Studies in Health Technology and Informatics*, pages 633–637, Sarajevo, Bosnia and Herzegovina, 2009. IOS Press.
- [9] SEYFANG, A., MIKSCH, S., MARCOS, M., WITTENBERG, J., POLO-CONDE, C., ROSENBRAND, K., Bridging the gap between informal and formal guideline representations. In *European Conference on Artificial Intelligence (ECAI-2006)*, volume 141 of *Frontiers in Artificial Intelligence and Applications*, pages 447–451, Riva del Garda, Italy, 2006. IOS Press.
- [10] SHIFFMAN, R.N., KARRAS, B.T., AGRAWAL, A., CHEN, R., MARENCO, L., NATH, S., GEM: a proposal for a more comprehensive guideline document model using XML. *Journal of the American Medical Informatics Association (JAMIA)*, 7(5):488–498, 2000.

Corresponding Author

Katharina Kaiser

Institute of Software Technology & Interactive Systems, Vienna University of Technology
Favoritenstrasse 9-11/188-1, 1040 Wien

Email: kaiser@ifs.tuwien.ac.at