

Semantische Annotierung von Arztbriefen zur Generierung diagnostischer Trainingsfälle

Alexander Hörnlein¹, Stanislaus Reimer², Christian Kneitz²,
Christian Betz¹, Frank Puppe¹

¹ Universität Würzburg,
Lehrstuhl für Künstliche Intelligenz und Angewandte Informatik
hoernlein | betz | puppe@informatik.uni-wuerzburg.de
² Poliklinik der Universitätsklinik Würzburg
Reimer_S | Kneitz_C@medizin.uni-wuerzburg.de

Abstract: Problembasiertes Lernen erfordert eine möglichst große Anzahl von Fallstudien. Jedoch ist die Erstellung von Trainingsfällen ein aufwändiger Prozess. Die Kerninformation ist in Fallbeschreibungen, z.B. Arztbriefen in der Medizin, bereits enthalten. Um daraus Trainingsfälle generieren zu können, ist eine Nachbearbeitung und semantische Annotierung erforderlich. Wir präsentieren einen neuartigen, skalierbaren Ansatz, der auf eine bequeme und inkrementelle Nachbearbeitung mit automatischer Fallgenerierung optimiert ist. Die Annotationen erfassen die Struktur der Dokumente in XML, das zur Ausführung in eine Wissensbasis des Autorensystems d3web.Train übersetzt wird. Erste Testergebnisse bewerten den Nutzen, auch von verschiedenem Zusatzwissen, aus Sicht der Endbenutzer (Medizinstudenten).

1 Einleitung

Beim problembasierten Lernen werden dem Benutzer Kasuistiken präsentiert, die dieser selbstständig löst und sich dabei kontextbezogen das nötige Hintergrundwissen aneignet [NS92]. Die kognitiven Fähigkeiten beim Lösen von Kasuistiken können gut am Computer trainiert werden, der einen Fall dem Benutzer schrittweise präsentiert, dazu Aufgaben stellt und Feedback zu den Antworten des Benutzers liefert [R00]. Solche fallbasierten Trainingssysteme sind inzwischen für viele medizinische Gebiete verfügbar [BFL02] [PAB03] [PBF04]. Die Erstellung der Trainingsfälle ist jedoch auch mit guten Autorenwerkzeugen (z.B. Seitenbasierten Autorensystemen wie [CASUS], datenbankbasierten Autorensystemen wie [CAMPUS], oder wissensbasierten Autorensystemen wie [d3web.Train]) eine aufwändige Aufgabe, die dadurch erschwert wird, dass sich der Autor in ein spezielles Autorenwerkzeug einarbeiten muss.

In diesem Beitrag stellen wir eine innovative Lösung vor, nämlich die Generierung von Trainingsfällen direkt aus Arztbriefen und demonstrieren ihre Umsetzung in einem ersten Prototyp mit d3web.Train (basierend auf in [RP99] sowie [BBH04] beschriebenen Vorarbeiten). Der Ansatz geht zurück auf [FD94]. Wir gehen darüber hinaus, indem wir die so erstellten Trainingsfälle durch Zusatzwissen schrittweise anreichern können, um den Trainingseffekt für den Benutzer zu verbessern, wobei das Zusatzwissen mit dem gleichen Textsystem eingegeben wird, mit dem der Arztbrief ge-

schrieben wurde. Die Generierung von Trainingsfällen aus Arztbriefen hat verschiedene Vorteile:

- Der Fallautor kann sehr schnell viele Fälle erstellen (falls gute Arztbriefe vorhanden sind) und nach Bedarf schrittweise verfeinern.
- Für den Benutzer kann das Fallangebot erheblich verbreitert werden.
- Für den Dozenten einer Universitätsklinik wird erstmalig die Möglichkeit geschaffen, dass er aktuelle Patienten seiner Station als fallbasierte Trainingssysteme in seinen Lehrveranstaltungen präsentiert, ohne sich mit Autorensystemen auskennen zu müssen. Im Vergleich zu fertig ausgearbeiteten, fremdbezogenen Trainingsfällen kann der Patient auch „live“ auftreten und der Dozent sein Hintergrundwissen über den Fall oder Zusatzinformationen nach Bedarf nutzen.

Im Folgenden definieren wir zunächst die Kerninformationen eines Trainingsfalls und beschreiben dann in Kap. 3 den Aufbau eines typischen Arztbriefes. Kap. 4 geht auf die notwendigen Annotationen und Transformationen ein. In Kap. 5 und 6 werden der Aufbau des resultierenden Trainingssystems und ein Beispielfall beschrieben. Kap. 7 beschreibt die Ergebnisse erster Tests mit Medizinstudenten. Kap. 8 schließt mit Diskussion und Ausblick.

2 Kerninformationen eines Trainingsfall

Wir haben in [BP03] die Kerninformationen eines diagnostischen Trainingsfalls (ohne Hintergrundwissen und freie Fragen) als 8-Tupel definiert:

$$F = (B, I, U, W_U, D, W_D, T, W_T)$$

- $B = \{B_i : i = 0 \dots b\}$ ist eine nichtleere Menge von Beobachtungen.
- $I = \{B_i : i = 0 \dots b'; b' \leq b\}$ ist eine initiale Menge von Beobachtungen.
- U, D, T sind drei disjunkte Mengen: Untersuchungen, Diagnosen und Therapien.
- W_U, W_D, W_T enthalten Wissen für Bewertungsfunktionen, die bei gegebener Menge von Beobachtungen angeben, welche Untersuchungen wie sinnvoll, welche Diagnosen wie wahrscheinlich und welche Therapien angemessen sind.
- Eine Untersuchung enthält eine Menge von Beobachtungen. Umgekehrt ist jede Beobachtung in mindestens einer Untersuchung enthalten.
- Beobachtungen können als Texte, Multimedia-Elemente, als kodierte Symptome oder als Kombination davon, ggf. mit Informationen über die Abbildung, repräsentiert sein.

Der Benutzer bekommt schrittweise Beobachtungen des Falles präsentiert und muss verschiedene Typen von (Auswahl)Aufgaben lösen:

- Untersuchungen indizieren. Das Feedback wird mit Hilfe des Untersuchungswissens W_U generiert.
- Diagnosen stellen. Feedback wird mit Hilfe des Diagnosewissens W_D generiert.
- Therapien angeben. Feedback wird mit Hilfe des Therapiewissens W_T generiert.

- Multimediale Beobachtungen befunden, d.h. als Text oder kodierte Symptome beschreiben. Das Feedback wird aus dem Abbildungswissen zwischen Multimedia-Elementen und kodierten Symptomen bzw. Text generiert.

Zunächst bekommt der Benutzer die initiale Menge von Beobachtungen, d.h. die Hauptbeschwerden des Patienten, präsentiert. Diese sind gewöhnlich in den Untersuchungen „Anamnese“ und „körperliche Untersuchung“ zusammengefasst. Danach kann der Benutzer weitere Untersuchungen indizieren (z.B. „EKG“, „Röntgen der Hände“, usw.), deren Beobachtungen ihm dann präsentiert werden, oder er kann vorläufige oder endgültige Diagnosen stellen. Wenn Beobachtungen nur als Multimedia-Elemente (z.B. EKG-Kurven, Röntgenbilder) präsentiert werden, müssen diese erst als Text bzw. kodierte Symptom befundet (z.B. „AV-Block I“ oder „Schenkelblock“) werden, bevor sie diagnostisch interpretiert werden. Weiterhin kann er Therapien angeben. Letztlich wird er danach beurteilt, ob er den virtuellen Patienten im Fall richtig diagnostiziert und therapiert hat. Außerdem kann man bewerten, wie hoch sein diagnostischer Aufwand (d.h. die Untersuchungskosten) war.

3 Aufbau eines Arztbriefes

Ein typischer Arztbrief (s. Abbildung 1) besteht aus einem für diagnostische Trainingsfälle nicht relevanten „Vorspann“ mit Absender, Adresse, Anrede usw. sowie einem „Nachspann“ mit Grüßen, Unterschrift usw. Der relevante Kern beginnt meist mit einer Liste der Diagnosen, enthält im Hauptteil die vollständige Liste der durchgeführten Untersuchungen jeweils mit ihren Ergebnissen (Beobachtungen) und am Ende findet sich die Beurteilung einschl. der Therapien und weiteren Maßnahmen. Daraus ergibt sich bereits eine grobe Zuordnung zu den Kerninformationen des Trainingsfalles:

- Die Diagnosen und die Untersuchungen des Arztbriefes lassen sich direkt auf die des Trainingsfalles abbilden. Die Beobachtungen, die jeweils zu einer Untersuchung gehören, liegen meist nur als ein zusammenfassender Text vor und weder als kodierte Symptome noch als Multimedia-Elemente.
- Die Therapien des Arztbriefes sind eher implizit in dem Beurteilungstext enthalten.
- Das Bewertungswissen ist in der Liste der Diagnosen und Untersuchungen (sowie implizit bei den Therapien) enthalten und gibt nur an, dass diese Objekte jeweils korrekt sind. Es gibt keine Informationen über Begründungen, über relative Gewichte und Wahrscheinlichkeiten oder über Alternativen.

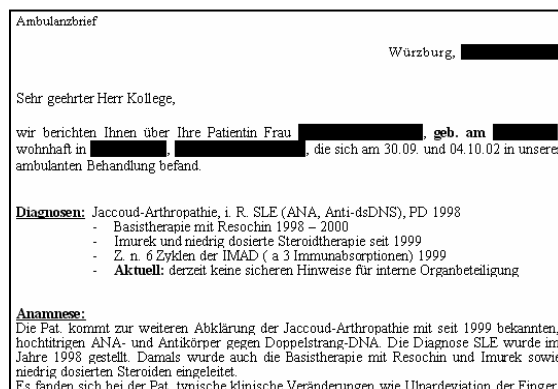


Abbildung 1: Beginn eines Arztbriefes mit Vorspann, Übersicht über Diagnosen und Therapien und Beobachtungen zur Anamnese

- Die initialen Beobachtungen sind nicht explizit aufgeführt, aber stehen häufig am Anfang der Kerninformationen.

Aus diesen Überlegungen ergibt sich, dass man aus Arztbriefen wichtige Informationen für einen Trainingsfall automatisch extrahieren kann (Diagnosen, Untersuchungen, Beobachtungen), teilweise mit Heuristiken (Initiale Beobachtungen). Das Bewertungswissen wird dabei so interpretiert, dass die angegebenen Diagnosen und Untersuchungen richtig und für die Lösung des Falles notwendig und alle anderen falsch sind. Im Folgenden beschreiben wir, wie mit wenig Aufwand der Arztbrief um weitere Informationen angereichert werden kann.

4 Annotation eines Arztbriefes

4.1 Ontologische Standardisierung

Bei der semantischen Anreicherung von Texten (z.B. für das Semantic Web) ist eine Festlegung auf eine standardisierte terminologische Ontologie sehr nützlich [FHH01]. Von den vier Bereichen (Diagnosen, Therapien, Untersuchungen, Beobachtungen) ist eine begriffliche Standardisierung in der Medizin bei Diagnosen gut (z.B. ICD-10), bei Therapien und Untersuchungen mäßig, und bei Beobachtungen eher schlecht gegeben. Allerdings muss diese in den Arztbriefen nicht eingehalten sein. Auch ist ein Standard wie der ICD-10 nicht für tutorielle Zwecke entworfen. Daher ist es vorteilhaft, wenn für die Annotation eines Arztbriefes zunächst eine Standardterminologie festgelegt wird, die die didaktisch wichtigsten Diagnosen, Therapien und Untersuchungen der Domäne enthält. Wenn auf Standards wie den ICD-10 zurückgegriffen wird, ist es meist sinnvoll, sich auf eine relevante Teilmenge davon zu beschränken. Eine hierarchische Strukturierung mit Ober- und Unterbegriffen ist für tutorielle Zwecke hilfreich.

Um einen Wiederverwendungseffekt zu erzielen, sollte die Standardterminologie weder zu allgemein (d.h. nicht auf die gesamte Medizin) noch zu speziell (d.h. nicht genau auf einen Arztbrief bezogen) sein, sondern auf das didaktische Fachgebiet zugeschnitten sein, aus dem der Arztbrief stammt (z.B. Rheumatologie, Kardiologie usw.). Sie ist dann für die Aufbereitung aller Arztbriefe dieses Fachgebietes anwendbar. Wenn im Arztbrief die Begriffe der Standardterminologie nicht benutzt sind, gibt es zwei Möglichkeiten, die Kongruenz herzustellen:

- Manuelles Editieren des Arztbriefes, indem die entsprechenden Begriffe ersetzt werden (z.B. mittels Drag-and-Drop aus der standardisierten Begriffshierarchie).
- Erweiterung der Standardterminologie um Synonymlisten, so dass eine automatische Zuordnung erfolgen kann.

4.2 Auszeichnung von trainingsfallrelevanten Informationen

Da im Arztbrief die Therapien nur implizit im Beurteilungstext enthalten sind, ist für die Bewertung des Trainingsfalls deren explizite Auflistung in der Standardterminologie erforderlich. Darüber hinaus sollte das Bewertungswissen verfeinert werden. Das umfasst:

- bei den Diagnosen eine Angabe, wie sicher die Diagnose ist (nicht unbedingt als numerische Wahrscheinlichkeit, sondern eher als Kategorie wie „verdächtig“ und „bestätigt“) und wie wichtig eine Diagnose in Relation zu anderen Diagnosen ist (da der Student gut bewertet werden soll, wenn er zwar nicht alle, aber die wichtigen Diagnosen des Falles erkennt).
- bei den Therapien ebenfalls die Angabe, wie wichtig sie in Relation zu anderen Therapien ist (z.B. wird häufig zwischen kausalen und symptomatischen Therapien unterschieden, wobei erstere die Ursachen der Krankheit, während letztere nur ihre Symptome behandeln). Außerdem gibt es häufig annähernd gleichwertige Therapiealternativen für eine Diagnose, die ebenfalls angegeben werden sollten. Auch bei Therapien lässt sich eine Einstufung bezüglich der Sicherheit in „indiziert“ und „sinnvoll“ vornehmen.
- bei den Untersuchungen Angaben über deren Nutzen, der sich im Allgemeinen aus der Überprüfung einer bestimmten Diagnose ergibt, sowie deren Kosten (finanziell, aber auch im Hinblick auf Risiken für den Patienten).

Während sich die Begründungen für Therapien und Untersuchungen meist aus den Diagnosen ergeben, sind für die Begründungen der Diagnosen die Beobachtungen relevant. Allerdings sollte zur Angabe von Begründungswissen für Diagnosen die Beobachtungsterminologie ebenfalls standardisiert werden, was wie oben angedeutet sehr aufwändig ist. Als einfachen Ersatz für regelbasiertes Bewertungswissen kann man für tutorielle Zwecke nach den wichtigen Untersuchungen die Zwischendiagnosen explizit angeben. Diese werden im Arztbrief nach den jeweiligen Untersuchungsergebnissen im gleichen Format wie die Hauptdiagnosen aufgelistet.

Ebenfalls sehr nützlich ist die Ergänzung des Arztbriefes um Multimedia-Elemente. Diese werden dann ebenfalls nach den Untersuchungen eingefügt, zu denen sie gehören (z.B. Röntgen-Bilder nach der Untersuchung „Röntgen des Thorax“).

4.3 Annotation in XML

Den Trainingsfall konvertieren wir zunächst in ein strukturiertes XML-Dokument über einer DTD (vgl. auch [MRS02]), die die Auszeichnungen im Arztbrief abbildet:

```

<!Element Arztbrief (Diagnose+, Therapie*, InitialeBeobachtungen,
                    Untersuchungsgruppe+, Schlussbetrachtung)>
<!Element Diagnose (#Pcdata)>
<!Attlist Diagnose      Gewicht (sehr wichtig | wichtig | relativ wichtig | normal) „normal“
                        Sicherheit (verdächtig | bestätigt) „bestätigt“>
<!Element Therapie (#Pcdata)>
<!Attlist Therapie      Gewicht (sehr wichtig | wichtig | relativ wichtig | normal) „normal“
                        Sicherheit (sinnvoll | indiziert) „indiziert“>
<!Element InitialeBeobachtungen (#Pcdata)>
<!Element Untersuchungsgruppe (Untersuchungsblock+ Diagnose+ Therapie*)>
<!Element Untersuchungsblock (Untersuchung Beobachtungen Multimedia*)>
<!Element Untersuchung (#Pcdata)>
<!Element Beobachtungen (#Pcdata)>
<!Element Multimedia (#Pcdata)> ;;; eine URL auf ein Multimedia-Element (z.B. ein Bild)
<!Element Schlussbetrachtung (#Pcdata)>

```

Zusätzlich geben wir eine DTD für die terminologische Ontologie der Diagnosen, Therapien und Untersuchungen an:

```
<!Element Ontologie (Diagnose*, Therapie*, Untersuchung*)>
<!Element Diagnose ((#Pcdata), Verweis?, Diagnose*)
<!Element Therapie (#Pcdata), Verweis?, Therapie*)
<!Element Untersuchung (#Pcdata), Verweis?, Untersuchung*)
<!Element Verweis (#Pcdata)> ;; eine URL auf Zusatzinformationen zum Objekt
```

Zwar können mit dieser Definition beliebig tiefe Hierarchien erzeugt werden, im Allgemeinen reichen aber wenige Ebenen von Diagnosen, Therapien und Untersuchungen aus. So sind Diagnosehierarchien wichtig, weil aufgrund der unvollständigen Symptomatik oft nur eine Grobdiagnose gestellt werden kann.

4.4 Transformationen

Die Umwandlung des Arztbriefes in das XML-Dokument erfolgt durch einen Parser aufgrund von möglichst bequem einzugebenden Textauszeichnungen. Dazu eignen sich Schlüsselwörter bzw. -zeichen, Absätze sowie Textauszeichnungen **Fett**, *Kursiv*, und Unterstrichen für die Attribute mit jeweils zur DTD passenden Semantik (s. Abbildung 2 links). Auch die Ontologie kann aus einer Textdatei eingelesen werden, wobei die Hierarchiestufen z.B. durch Einrückungen mit Sonderzeichen gekennzeichnet sind.

Schließlich kann das XML-Dokument leicht in eine Wissensbasis im Format von d3web (ebenfalls eine XML-Struktur) transformiert werden. Für die Autorenkomponekte von d3web gibt es zwar spezielle Editoren, mit denen außer der hier dargestellten noch eine wesentlich mächtigere Wissensrepräsentation unterstützt wird, aber die Benutzung dieser Editoren ist für die Bearbeitung von Arztbriefen nicht erforderlich.

5 Darstellung des Trainingssystems aus Benutzersicht

Die (webbasierte) Oberfläche behandelt Fälle, in denen die Informationen kodiert oder als Freitext bzw. als Mischform vorliegen, im Wesentlichen gleich. Der Lerner hat folgende Interaktionsmöglichkeiten:

- Anfordern von Untersuchungen
- Stellen von Diagnosen
- Wählen von Therapien
- Anzeigen und ggf. Befunden von zusätzlichen Multimedia-Daten (derzeit hauptsächlich Bilder)
- Anzeigen von allgemeinen Informationen (Hintergrundwissen) sowie einem fallbezogenen Abschlusskommentar
- Hilfe und Feedback

Je nachdem, welches Wissen in dem Fall vorliegt, werden Interaktionsmöglichkeiten, für die die nötigen Informationen fehlen, ausgeblendet oder eingeschränkt und die Bewertung entsprechend angepasst. Bei Arztbrief-basierten Fällen liegt typischerweise zunächst nur Wissen zu Anforderung von Untersuchungen und zum Stellen von Diagnosen und Therapien vor.

Der Lerner erhält vor dem eigentlichen Beginn des Falls die Einführung in den Fall (Intro), anhand derer er entscheiden kann, ob er den Fall tatsächlich bearbeiten will

oder nicht. Es gibt zwei Möglichkeiten der Bearbeitung: Der geführte Modus und der freie Modus, die sich vor allem hinsichtlich der Untersuchungsauswahl und des Feedbacks zu Arbeitsdiagnosen unterscheiden:

- Beim geführten Modus wird die Anforderung von Untersuchungen vom System gruppenweise in einer festen Reihenfolge vorgenommen. Eine Bewertung der Untersuchungsauswahl des Lerner fällt somit. Dafür können nach jeder Anforderung eines Untersuchungsblocks die Arbeitsdiagnosen vom Lerner erfragt und mit den im Arztbrief angegebenen Zwischendiagnosen verglichen werden.
- Beim freien Modus muss der Lerner hingegen selbst die nötigen Untersuchungen auswählen. Dabei kann er völlig frei Untersuchungen in beliebiger Reihenfolge anfordern. Arbeitshypothesen können jedoch nur entweder mit den Abschlussdiagnosen oder mit den Zwischendiagnosen nach Abarbeitung kompletter Untersuchungsblöcke verglichen werden.

Das Ergebnis der Untersuchungen wird bei kodierten Informationen als tabellenartige Liste, bei Freitext-Daten als fortlaufender Text angezeigt. Wenn Formatierungsanweisungen als HTML-Code extrahiert werden können, dann werden diese Formatierungen entsprechend umgesetzt.

Bei der Eingabe von Diagnosen, Therapien und Untersuchungen wählt der Benutzer aus der Hierarchie aller in der Terminologie vorgesehenen Objekte aus (vergleichbar mit MC-Fragen im Long-Menu-Format [SVS96]).

Der Lerner kann den Fall jederzeit abschließen - er erhält dann eine Bewertung, deren Umfang sich wie beim Feedback nach den vorhandenen Informationen richtet. Nach der Bewertung erhält er den abschließenden Text des Autors, der bei generierten Fällen aus „Beurteilung und Verlauf“ extrahiert wird.

6 Beispiel

Wir zeigen nun anhand eines Abschnitts in einem Arztbrief, wie sich ein generierter Fall im Trainingssystem für den Lerner im geführten Modus darstellt (s. Abbildung 2).

Der Abschnitt aus dem Arztbrief beschreibt die erste angegebene Untersuchungsgruppe. Deren Anfang ergibt sich dadurch, dass die enthaltenen Untersuchungen die ersten des Arztbriefes sind, beendet wird sie von einem Diagnoseblock. Startet der Lerner einen Fall im geführten Modus, dann werden die Untersuchungen der ersten Untersuchungsgruppe sofort angefordert, zunächst bekommt der Lerner aber den kursiv ausgezeichneten Abschnitt aus den Beobachtungen der ersten Untersuchung (hier: „Anamnese“) als Fall-Intro präsentiert. Erst wenn er den Abschnitt „Anamnese“ wählt, erhält er den vollen Text der zur zugehörigen Untersuchung angegeben ist. Anschließend sollte der Lerner die weiteren Ergebnisse aus dem Bereich „Körperlicher Untersuchungsbefund“ anzeigen lassen. Damit erhält er auch Zugriff auf die vier im Arztbrief eingefügten Bilder zu dieser Untersuchung. Diese werden übersichtlich in einem neuen Fenster angezeigt.

Wenn der Lerner nun weitere Untersuchungen anfordern will, wird er darauf aufmerksam gemacht, dass er vorher noch Verdachtsdiagnosen angeben sollte. Dabei wird ihm vorgegeben, wie viele Diagnosen (auf welchen Ebenen der Hierarchie) mit welcher Bewertung („verdächtig“, „bestätigt“) erwartet werden. Hat er Diagnosen eingegeben und fordert die nächsten Untersuchungen an, dann wird ein Feedback basierend auf den Angaben im zugehörigen Diagnoseblock generiert. Der Autor kann

diese Diagnosen durch Kombinationen von Markierungen (normal, fett, unterstrichen) gewichten. Kursiv markierte Diagnosen werden als „verdächtig“ eingestuft, andernfalls als „bestätigt“.

Zusätzlich kann zu allen Begriffen der Ontologie Hintergrundwissen bereitgestellt werden. Im Gegensatz zur Evaluation (Kap. 7) ist in diesem Beispiel zwar keines vorhanden, es wird allerdings zu jedem Begriff ein Link auf eine entsprechende Suchanfrage generiert (im Beispiel bei www.google.de).

Danach werden alle Untersuchungen bis zum nächsten Diagnoseblock angezeigt – ist keiner mehr vorhanden, dann werden die restlichen angezeigt und die erwarteten Diagnosen richten sich nach den am Beginn des Arztbriefs angegebenen Diagnosen, die sich auf den kompletten Fall beziehen. Typischerweise sind in Zwischen-Diagnoseblöcken nur Diagnosen mit der Bewertung „verdächtig“ angegeben; in einem solchen Fall kann der Lerner erst am Ende des Falles Therapien wählen. Es kann aber auch vorkommen, dass eine Diagnose bereits als „bestätigt“ gewertet wird und zusätzlich eine Therapie als „indiziert“ angegeben ist. Dies soll eine Sofortmaßnahme widerspiegeln, die während der Untersuchung durchgeführt wurde und hier muss der Lerner bereits vor Ende des Falles Therapien wählen.

Anamnese:

32-jährige Pat. stellt sich mit ausgeprägten Deformitäten im Bereich der Hände in ihrer Sprechstunde vor. Die Fingerdeformitäten sind der Pat. vor ca. einem Jahr aufgefallen. Seit ca. 6 Monaten hat die Pat. Probleme beim Erledigen von Tipparbeit im Büro.

Es finden sich bei der Pat. klinische Veränderungen wie Ulnardeviation der Finger, symmetrische Subluxationen der Metakarpophalangealgelenke, leichte Hyperextension der Finger DIG IV und V bds., Streckdefizit der DIG II-V, re. etwas stärker ausgeprägt als li. sowie Schwannenhalsdeformationen ohne das Vorhandensein von klinischen Arthritiszeichen, wie Gelenkschwellungen und Druckschmerzhaftigkeit.

Die Pat. verneint: Fieber, Morgensteifigkeit, Sicca- und Raynaud-Symptomatik, Augen- und Hamwegentzündungen, Stuhlrhythmusstörungen.

Gewichtsschwankungen bis zu 3-4 kg im Jahr. Das Hautproblem der Pat. sind die Deformitäten an den Händen und Hautveränderungen im Gesicht- und Halsbereich, sonst sei die Pat. beschwerdefrei. Auf die Frage bezüglich Allergien, antwortet die Pat., dass sie eine Sonnenallergie hätte.

Körperlicher Untersuchungsbefund:

32-jährige Pat. in gutem AZ und EZ (175 cm/64 kg). Untersuchung von Kopf und Hals unauffällig. Keine peripheren Lymphknotenschwellungen tastbar. Vesikuläres Atemgeräusch mit guter Zwerchfellverschieblichkeit. Herzaktion regelmäßig mit 71/Min. Blutdruck 120/80 mmHg. Herzton rein. Abdominelle Untersuchung unauffällig.

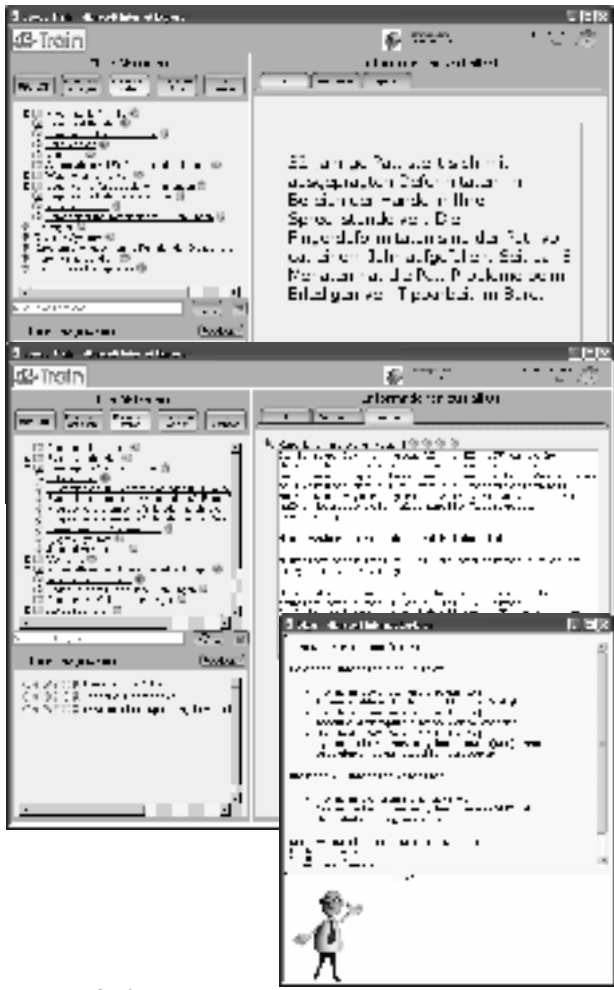
Hauterythem im Gesicht- und Halsbereich.

Neurologischer Status: Reflexe symmetrisch auslösbar, insgesamt unauffällig.

Gelenkstatus: Gaendlen-Zeichen bds. negativ, keine druckschmerzhaften Gelenke, keine direkten Arthritiszeichen, keine Schwellungen, Ulnardeviation der Finger bds., Subluxationen der MCP und PIP-Gelenke, Schwannenhalsdeformationen. Die Untersuchung der übrigen peripheren großen Gelenke und Füße erbrachte keinen pathologischen Befund (Z. n. Hallux valgus-OP bds.).



Zusatzdiagnosen:
 Es sind Arthritiszeichen bds. vorhanden, jedoch ohne Gelenkschwellungen und Druckschmerzhaftigkeit.
 Die Untersuchung der übrigen peripheren großen Gelenke und Füße erbrachte keinen pathologischen Befund (Z. n. Hallux valgus-OP bds.).



254

Abbildung 2: Links: Ausschnitt aus einem annotierten Arztbrief. Rechts: Der entsprechende Abschnitt aus dem Ablauf des generierten Trainingsfalles. Oben: Beginn des Falles mit Fall-Intro, Mitte: Ergebnisse der Untersuchungsgruppe „Körperlicher Untersuchungsbefund“ Unten: Erzeugtes Feedback (Vollansicht)



Abbildung 3: Trainingssystem im freien Modus. Hintergrund: Freie Auswahl der Untersuchungen auf der linken Seite, tabellarische Auflistung der Ergebnisse auf der rechten Seite, ganz rechts ist der Kostenbalken zu sehen. Vordergrund: Bildbefundung, links die Auswahl der möglichen Ergebnisse und befundbaren Untersuchungen, rechts Überblick über alle verfügbaren Daten mit ausgewähltem Bild.

Aus Arztbriefen lassen sich auch Fälle für den freien Modus generieren, allerdings ohne Feedback für Zwischendiagnosen. Dieses erfordert regelbasiertes Zusatzwissen. In solchen wissensbasierten Fällen sind andere Features möglich, die auch evaluiert wurden:

- *Tabellarische Auflistung der Symptome:* Im Gegensatz zum Fließtext werden Ergebnisse der Untersuchungen tabellarisch aufgeführt (s. Abbildung 3). Dabei kann zu jedem Einzelergebnis zusätzliches Hintergrundwissen hinterlegt sein. Die dazu nötigen Daten lassen sich weder einfach im Arztbrief überblickbar angeben noch extrahieren – dazu wäre die Erstellung einer sehr umfangreichen Ontologie nötig, die wiederum eine entsprechend gründliche Überarbeitung aller Arztbriefe bedingt.
- *Feedback für Zwischendiagnosen bei freier Untersuchungsauswahl:* Im freien Modus kann der Lerner aus der Untersuchungshierarchie in beliebiger Reihenfolge Untersuchungen auswählen (s. Abbildung 3), deren Ergebnisse ihm präsentiert

- werden. Nur bei der freien Untersuchungsauswahl ist eine Kostenbeachtung sinnvoll. Zusätzlich kann er jederzeit Zwischendiagnosen angeben und sich ein Feedback relativ zu den ihm bis dahin bekannten Symptomen anzeigen lassen.
- *Bildbefundung*: Bei der Bildbefundung wird dem Lerner das Ergebnis einer Untersuchung nicht mitgeteilt, sondern er wird dazu aufgefordert, diese Befundung selbst vorzunehmen. Dazu erhält er Zugriff auf alle zum Fall verfügbaren Multimedia-Daten und muss aus den möglichen Ergebnissen dieser Untersuchung eines oder mehrere auswählen (s. Abbildung 3). Auch solche Zusammenhänge können bisher nur mit dem Autorenwerkzeug erstellt werden. Die bloße Bilddarstellung findet ebenfalls in diesem Fenster statt, der linke Teil (Untersuchungsauswahl und -beantwortung) wird dabei geeignet ersetzt.

7 Erste Evaluationsergebnisse

Zunächst wurde eine entwicklungsbegleitende Vorstudie mit fünf kurz vor dem dritten klinischen Staatsexamen stehenden Studenten der Poliklinik der Universitätsklinik Würzburg durchgeführt, bei der anhand von sieben Kasuistiken das generelle Systemdesign und insbesondere verschiedene Merkmale des Trainingssystems getestet wurden. Bei dem Fragebogen (s.u.) zeigte sich eine relativ hohe Akzeptanz. Bezüglich der Merkmale des Trainingssystems legten die Studenten viel Wert auf Bildmaterial, Hintergrundwissen sowie einer Abschlussdiskussion des Falles und bevorzugten echte Fälle gegenüber Musterfällen, während die Präferenzunterschiede zwischen Arztbrief-generierten und Wissensbasis-generierten Fällen sowie zwischen dem freien und dem geführten Modus weniger eindeutig waren. Außerdem zeigten sich Probleme bei technischen Aspekten, die mit Noten zwischen 2,4 und 2,8 deutliche schlechtere Bewertungen als die Gesamtbeurteilung hatte (1,8).

Ziel der anschließenden Evaluation (vgl. auch Evaluation des Vorgängersystems in [RKT04]) war es, die generelle Akzeptanz des Systems bei Studenten in einem Pflichtkurs (Hauptvorlesung der Inneren Medizin, beschränkt auf den Teilbereich Rheumatologie) sowohl an sich als auch als Ergänzung zur Vorlesung zu erfassen. Die Nutzung des Systems war freiwillig. Es wurde während der Vorlesung mit ca. 130 Medizinstudenten im 6.-8. Semester zehn Minuten vorgeführt und dann den Studenten auf einem Handzettel eine kurze Anleitung zur Bedienung sowie persönliche, anonyme Zugangsdaten mitgeteilt. Für jede in der Vorlesung behandelte Diagnosenfamilie wurden den Studenten zwei Fälle angeboten, wobei zum Zeitpunkt dieser Zwischenauswertung (14.06.04) nur acht Fälle von insgesamt zwölf für die Studenten freigeschaltet waren und in die Evaluation eingehen konnten. Die Studenten bearbeiteten die Fälle von zu Hause aus über das Internet ohne Hilfestellung durch einen Tutor.

Insgesamt wurden zum Zeitpunkt dieser Zwischenauswertung von 36 Studenten 142 Fälle mit einer durchschnittlichen Zeitdauer von 12 ± 9 Minuten bearbeitet. Nach Bearbeitung jedes Falles wurden zwei Fragebögen angeboten, einer mit 9 Fragen zum Trainingssystem als Ganzem und einer mit 3 Fragen zum Fall. Dabei wurden 55 Fallfragebögen von 19 Studenten beantwortet, sowie der Systemfragebogen von 15 verschiedenen Studenten (von manchen auch mehrfach, wobei nur der zuletzt beantwortete Fragebogen in die Auswertung einging).

Die Fragen wurden mit Schulnoten von 1 - 5 (1 = sehr gut, 5 = sehr schlecht) bzw. bei den Fragen zur Gesamtnote mit Schulnoten von 1 - 6 bewertet. Die allgemeinen Fragen bezogen sich auf:

- technische Aspekte (3 Fragen): Einarbeitung, Optik, Bedienung
- Nützlichkeitsaspekte (3 Fragen): Erwerben neuer Kenntnisse, zukünftige Nutzung des Systems und nützliche Ergänzung zu anderen Lernformen
- Gesamtbewertung (3 Fragen): Gesamtnote System, Gesamtnote System als Ergänzung zur Vorlesung, Wunsch nach weiterer Nutzung in anderen Lehrveranstaltungen.

Die drei fallspezifischen Fragen bezogen sich auf die Eigenschaften des Falles: wie lehrreich, angemessener Schwierigkeitsgrad und angemessenes Feedback

Die Beantwortung der Fragen zeigte generell eine hohe Zufriedenheit mit dem Trainingssystem und den Fällen. Die Durchschnittsnoten der 15 Systemfragebögen sind:

- Einarbeitung: $1,9 \pm 0,8$
- Optik: $1,5 \pm 0,6$
- Bedienung: $2,4 \pm 1,1$
- Erwerben neuer Kenntnisse: $2,0 \pm 1,0$
- Zukünftige Nutzung: $1,3 \pm 0,5$
- Nützliche Ergänzung zu anderen Lehrformen: $1,3 \pm 0,6$
- Gesamtnote: $1,5 \pm 0,5$
- Gesamtnote als Ergänzung zur Vorlesung: $1,5 \pm 0,9$
- Wunsch nach Nutzung in anderen Lehrveranstaltungen: $1,9 \pm 0,4$

Die Durchschnittsnoten aus den 55 ausgefüllten Fallfragebögen waren:

- Wie lehrreich war der Fall? $1,5 \pm 0,7$
- War der Schwierigkeitsgrad angemessen? $1,3 \pm 0,6$
- War das Systemfeedback angemessen?
fair: 82,6%; zu gut: 0%; zu schlecht: 13%; unpassend: 2,2%; k.A.: 2,2%

11 der 15 Studenten gaben auch Freitextkommentare ein, die fast durchweg positiv waren (z.B. „Tolles Programm, sollte auch bei anderen Veranstaltungen verwendet werden. Danke für die Mühe!“). Als wichtigster Wunsch wurde eine CD als Ergänzung zum Internet-Einsatz gefordert; ein Student monierte den unübersichtlichen Bildschirmaufbau.

8 Diskussion und Ausblick

Auch wenn die Evaluationsergebnisse eher Zwischenergebnisse einer laufenden Studie sind und wegen der Freiwilligkeit der Teilnahme nur begrenzte statistische Aussagen zulassen, deuten sie darauf hin, dass die Studenten das Bearbeiten von Trainingsfällen am Computer generell gut akzeptieren (was in Übereinstimmung mit vielen anderen Studien steht). Das gilt insbesondere auch für die aus Arztbriefen generierten Trainingsfälle. In der Vorstudie zeigte sich deutlich die Bedeutung des Zusatzwissens in den Fällen (fallbezogene Bilder, allgemeines Hintergrundwissen, fallbezogener Abschlusskommentar), was in der Evaluationsstudie berücksichtigt wurde. Das zeigt die Notwendigkeit einer entsprechenden Nachbereitung der Arztbriefe an.

Wir erwarten, dass die direkte Umsetzung von Arztbriefen in fallbasierte Trainingssysteme deren Akzeptanz in der medizinischen Ausbildung vor allem bei den Dozenten beträchtlich steigert, da diese ihre eigenen Fälle ohne großen Aufwand nutzen können. Natürlich sind weitere Evaluationen, insbesondere auch in Pflichtkursen, notwendig. Von besonderer Bedeutung wird sein, wenn die regulären Prüfungen auf Kasuistiken umgestellt werden.

Literaturverzeichnis

- [BBH04] Betz, C., Buscher, H.-P., Hörnlein, A., Puppe, F., Schuhmann, M.: Generierung diagnostischer Trainingsfälle aus Arztbriefen, in [P 04], 25-36, 2004.
- [BFL02] Bernauer, J., Fischer, M., Leven, J., Puppe, F., Weber, M. (Hrsg.): Rechnergestützte Lehr- und Lernsysteme in der Medizin: Proc. 6. Workshop der GMDS AG Computergestützte Lehr- und Lernsysteme in der Medizin, Shaker, Aachen, 2002.
- [BP03] Betz, C. und Puppe, F.: Formale Beschreibung fallbasierter diagnostischer Trainingssysteme, in [Puppe et al. 03], 12-23, 2003.
- [CAMPUS] Computergestützte Aus- und Weiterbildung in der Medizin mit dem flexiblen und simulativen fallbasierten System CAMPUS. <http://campus.fh-heilbronn.de> (14.06.04).
- [CASUS] Fallorientiertes Lernen in der Medizin. <http://link.medinn.med.uni-muenchen.de/instruct> (14.06.04).
- [d3web.Train] Fallbasierte intelligente Trainingssysteme für die Diagnostik. <http://www.d3webtrain.de> (14.06.04).
- [FD94] Felciano, R.M. and Dev, P. Multimedia clinical simulation based on patient records: Authoring, user interface, pedagogy. Proceedings of the 18. Annual Symposium on Computer Applications in Medical Care. 1994:59-63.
- [FHH01] Fensel, D., Horrocks, I., van Harmelen, F., McGuinness, D., and Patel-Schneider, P. F.: OIL: Ontology Infrastructure to Enable the Semantic Web, IEEE Intelligent System, 16(2), 2001.
- [MRS02] Merz, A.-K., Rockmann, F., Schwarz, C. und Reng, M.: MedicCaseML: XML Austauschformat für CBT Systeme in der medizinischen Aus- und Weiterbildung. In: [Bernauer. et al. 02]
- [NS92] Norman, G. R., Schmidt, H. G.: The psychological basis of problem-based learning – a review of the evidence. In: Academic Medicine 67 (9), 557-565, 1992
- [PAB03] Puppe, F., Albert, J., Bernauer, J., Fischer, M., Klar, R., Leven, J. (Hrsg.): Rechnergestützte Lehr- und Lernsysteme in der Medizin: Proc. 7. Workshop der GMDS AG Computergestützte Lehr- und Lernsysteme in der Medizin, Shaker, Aachen, 2003.
- [PBF04] Pöpl, S., Bernauer, J., Fischer, M., Handels, H., Klar, R., Leven, J., Puppe, F., Spitzer, K. (Hrsg.): Rechnergestützte Lehr- und Lernsysteme in der Medizin: Proc. 8. Workshop der GMDS AG Computergestützte Lehr- und Lernsysteme in der Medizin, Shaker, Aachen, 2004.
- [R00] Reinhardt, B.: Didaktische Strategien in generierten Trainingssystemen zum diagnostischen Problemlösen, Dissertation an der Universität Würzburg, infix-Verlag, disk 234, 2000.
- [RKT04] Reimer, S., Kneitz, C., Tony, H.-P., Schewe, S., Hörnlein, A., Puppe, F.: d3web.Train: Erste Evaluationsergebnisse zum Einsatz in der Mediziner Ausbildung an der Medizinischen Poliklinik der Universität Würzburg, in: [Pöpl et al. 04], 155-164, 2004.
- [RP99] Reinhardt, B. und Puppe, F.: Generierung flexibler intelligenter Trainingssysteme in: Künstliche Intelligenz Heft 4/1999, 12-18, 1999.
- [SVS96] Schuwirth, L. W. T., Van der Vleuten, C.P.M., Stoffers, H.E.J.H., & Peperkamp, A.G.W.: Computerized long-menu questions as an alternative to open-ended questions in computerized assessment. In: Medical Education 30, 50-55, 1996.