

Evaluation eines adaptiven Hilfesystems für Bedienprobleme mit d3web.Train

Alexander Hörnlein, Frank Puppe

Lehrstuhl für Künstliche Intelligenz und Angewandte Informatik der Universität Würzburg
{hoernlein, puppe}@informatik.uni-wuerzburg.de

Abstract

Simulative fallbasierte Trainingssysteme in der Medizin sollen es den Lernern ermöglichen, einen virtuellen Patienten so zu untersuchen und zu diagnostizieren wie sie es später bei einem realen Patienten können sollten. Um dem gerecht werden zu können, benötigen diese Systeme eine komplexere Oberfläche, die es wiederum dem Lerner schwieriger macht, den Umgang mit einem solchen System zu lernen. Deshalb haben wir für das Trainingssystem d3web.Train ein adaptives Hilfesystem auf der Basis eines Overlaymodells entwickelt, das den Lerner auf Bedienfehler aufmerksam macht. In diesem Papier präsentieren wir Ergebnisse einer Evaluation des Einsatzes im Wintersemester 2004/2005 im Kurs der Rheumatologie der Medizinischen Poliklinik der Universität Würzburg.

1 Einleitung

d3web.Train [d3web.Train, HBP02] ist ein intelligentes fallbasiertes Trainingssystem im WWW. Der Benutzer soll sich in dem System wie ein Arzt verhalten, der einen Patienten über die elektronische Patientenakte (EPA) ferndiagnostiziert. Dazu bietet d3web.Train dem Lerner so viele Handlungsmöglichkeiten wie sinnvoll und bei einer Webapplikation machbar. Dadurch wird aber die Oberfläche (siehe Abb. 1) komplexer als bei einem System, das sich z.B. auf Multiple-Choice-Fragen beschränkt [SHPM05]. Diese Komplexität ist gerade bei einem Trainingssystem besonders problematisch, da der Benutzer im Trainingskontext seine Konzentration schon benötigt, um die Inhalte, die ihm durch die Arbeit mit dem System vermittelt werden sollen, aufzunehmen. Bei bereits stattgefundenen Evaluationen [RKT04, HRKBP04, RHTKBPK05] wurde, obwohl das System insgesamt gute Noten bekam, immer wieder die Oberflächenergonomie kritisiert. Einige Studenten hatten sogar nach eigenen Angaben nur deshalb schlechte Ergebnisse, weil sie Teilaufgaben nicht gefunden und deshalb nicht bearbeitet hatten.

Um die Lerner auf d3web.Train vorzubereiten, bieten wir ihnen verschiedene herkömmliche Hilfen, unter anderem eine eingebaute kontext-abhängige Hilfe, vorgefertigte nach HTML exportierte Powerpoint-Präsentationen und eine kurze Einführung durch den Dozenten. Diese Angebote können aber folgende Probleme nicht lösen:

- Die Hilfe ist dem Kenntnisstand des Lerners nicht angepasst.
- Der Lerner weiß vielleicht gar nicht, dass er Hilfe benötigt

Die Klinik verfügt zwar über ein Skills Lab [Skills Lab], in dem die Studierenden individuelle Hilfe und Beratung erhalten können an. Dies hat aber den Nachteil, dass die Studierenden die Fallbearbeitung unterbrechen und das Skills Lab aufsuchen müssen.

Deshalb haben wir ein adaptives Hilfe-System implementiert, welches aufgrund der Aktionen des Lerners eine Einschätzung seiner Fähigkeiten vornimmt und durch geeignete Hilfetexte interveniert, wenn es feststellt, dass der Lerner Hilfe benötigt.

Dieses Papier ist folgendermaßen aufgebaut: in Kapitel 2 wird das Trainingssystem d3web.Train und das darin eingebettete Hilfesystem beschrieben, Kapitel 3 präsentiert Design, Ergebnisse und Diskussion der durchgeführten Evaluation und Kapitel 4 schließt mit Zusammenfassung und Ausblick.

2 d3web.Train und das adaptive Hilfesystem

Abhängig von den in einem Fall verfügbaren Daten und der Konfiguration durch den Autor kann bzw. muss der Lerner in einem Fall vier verschiedene Aufgabenarten bearbeiten:

- Untersuchungen anordnen: Dabei muss der Lerner entscheiden, welche der momentan sinnvollen Untersuchungen das beste Verhältnis zwischen Informationsgewinn und Kosten, Risiken, Zeitaufwand hat.
- Ergebnisse auswerten: Falls ein Untersuchungsergebnis in Form von Bildern oder anderen Multimedia-Elementen vorliegt (z.B. ein Röntgenbild), muss der Lerner die Bilder durch Beantwortung von MC-Fragen interpretieren.
- Diagnosen stellen: Basierend auf den Untersuchungsergebnissen muss der Lerner entscheiden, welche Diagnosen verdächtig werden können und welche bestätigt sind.
- Therapien festlegen: Sobald Diagnosen gestellt wurden muss der Lerner sinnvolle und indizierte Therapien auswählen.

Nach der Therapiefestlegung kann der Fall mit den Auswirkungen der Therapie zur weiteren Behandlung dem Lerner in einer Folgesitzung wieder vorgelegt werden.

Jede dieser Aufgaben besteht aus entweder teilweise geordneten atomaren Aktionen oder (Unter-)Aufgaben. Da die Oberfläche von d3web.Train auf HTML basiert, ist eine atomare Aktion entweder ein Klick oder das Abschicken eines Formulars. Die Architektur von d3web.Train folgt dem MVC-Paradigma [R79], wobei jede atomare Aktion durch den Lerner auf eine interne mit Parametern und Zeitstempel versehene Aktion abgebildet wird.

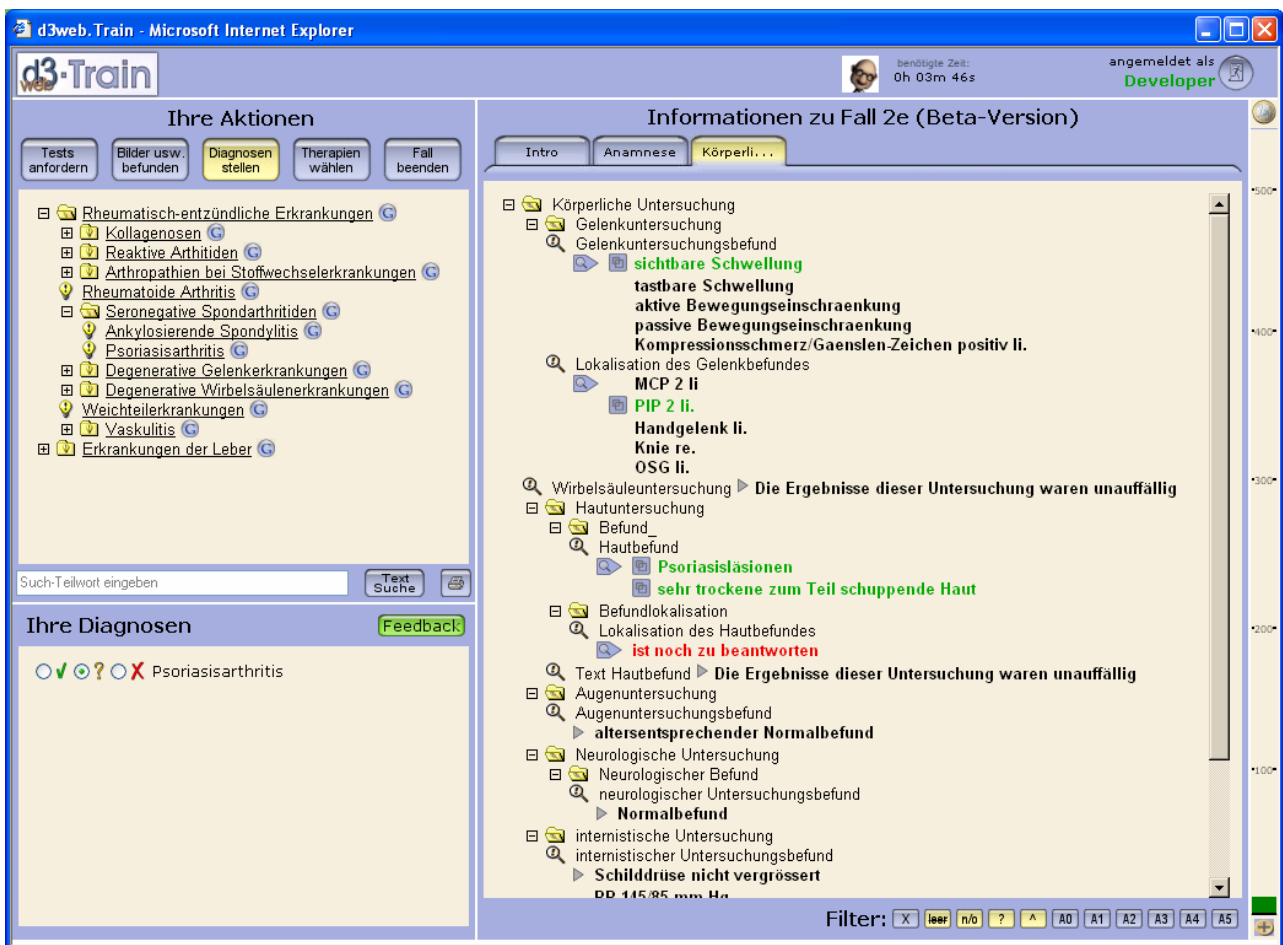


Abbildung 1: Oberfläche von d3web.Train Diagnose-Modus.

Nach [N01] ist eine wichtige Voraussetzung für eine Benutzermodellierung, dass das mentale Modell des Lernalers strukturell mit dem Experten-Modell des Systementwicklers übereinstimmt – andernfalls kann aus den Aktionen des Lernalers kein korrektes Modell seiner Fähigkeiten abgeleitet werden. Damit der Lerner also das System auf die richtige Weise wahrnimmt, versuchen wir durch eine einführende Darstellung der System-Metapher seine Sichtweise auf das System in die richtige Richtung zu lenken:

Stellen Sie sich vor, Sie sind Arzt, der eine Konsultation anhand einer elektronischen Patientenakte durchführt, die zunächst nur die Vorgeschichte enthält und schrittweise um weitere Untersuchungsergebnisse erweitert wird. Ihre Aufgabe ist das Stellen von Diagnosen, einschließlich der Befundung bildhafter Untersuchungsergebnisse.

Die elektronische Patientenakte im rechten Teil des Fensters ist in Abschnitte unterteilt, die durch Klick auf den zugehörigen Reiter (z.B. Anamnese, körperliche Untersuchung, ...) angezeigt werden.

Ihre Aufgabe ist es, aufgrund der Untersuchungsergebnisse verdächtige und schließlich bestätigte Diagnosen herzuleiten. Ihre jeweiligen Zwischendiagnosen wählen Sie in dem Baum im linken Fenster aus. Diese werden dann unten in einem immer sichtbaren Merktzettel angezeigt. Dort müssen Sie diese noch als "bestätigt" oder "verdächtig" bewerten.

Als Orientierung für die Anzahl der auszuwählenden Diagnosen finden Sie die gleiche Anzahl von Icons oberhalb des Merktzettels.

Nach der Diagnoseauswahl klicken Sie auf den Aktionsbutton "weiter" links oben. Dann werden neue Untersuchungsergebnisse in ihre Patientenakte eingetragen, aufgrund derer Sie Ihre Diagnosen auf dem Merktzettel anpassen sollen.

Abbildung 2: Erläuterung der Systemmetapher

Benutzermodellierung

Adaptive, also sich an den Benutzer anpassende Hilfe ist nur mit einem wissensbasierten System möglich, das die für die Hilfe wichtigen Fähigkeiten des Lernalers modelliert [FLS85]. Für unser Hilfesystem (ausführlicher in [HP04] beschrieben) haben wir dazu Overlay-Modellierung [CG77] gewählt, mit der wir die Kenntnisse des Lernalers über Konzepte bezüglich der Aufgaben und Aktionen bewerten. Diese Konzepte lassen sich aufteilen in solche über den richtigen Zeitpunkt für die Durchführung einer Aufgabe und solche über die richtige Durchführung einer Aufgabe selbst.

Der Kenntnisstand des Lernalers bzgl. Konzepten über den richtigen Zeitpunkt lässt sich leicht mit Regeln erfassen wie

- „Falls dem Lerner eine Untersuchung mit dem Text „ist noch zu bearbeiten“ präsentiert wird und der Lerner startet diese Bearbeitung (durch Klicken auf den Text) nicht, dann versteht er nicht, wann er „Ergebnisse auswerten“ muss.“
- „Falls der Lerner Untersuchungen anfordert ohne vorher Diagnosen angegeben zu haben, dann versteht er nicht, wann er „Diagnosen stellen“ muss.“

Es ist aber schwieriger festzustellen, ob ein Lerner weiß, wie er eine Aufgabe durchführen kann. Wenn der Lerner z.B. nicht weiß, dass er spezifischere Diagnosen nur auswählen kann, wenn er den richtigen Diagnose-Teil-Baum öffnet, dann würde er nur die Grob-Diagnosen der obersten Ebene auswählen. Wählt ein Lerner tatsächlich nur Grob-Diagnosen, dann lässt sich jedoch umgekehrt daraus nicht schließen, dass er die Navigation im Diagnose-Baum nicht versteht, denn er könnte (fälschlicherweise) auch ausschließlich Grob-Diagnosen als verdächtig oder bestätigt erkennen. In diesem Fall behelfen wir uns damit, dass wir dem Lerner angeben, wie viele Diagnosen welcher Ebene von ihm erwartet werden, so dass wir dann auch Regeln verwenden können wie

- „Falls der Lerner keine Fein-Diagnose eingibt, obwohl er Fein-Diagnosen eingeben muss, dann weiß er nicht, wie man zu einer Fein-Diagnose gelangt.“

Um den Kenntnisstand bzgl. eines Konzepts aus dem bzgl. eines diesem Konzept ähnlichen Konzepts abzuleiten, verwenden wir Regeln wie

- „Falls der Lerner weiß, wie er Fein-Diagnosen auswählt, dann weiß er wahrscheinlich auch, wie man spezifischere Therapien auswählt.“

Interventionen

Nach jeder Aktion des Benutzers wird das Benutzermodell aktualisiert, dabei werden zuerst alle Regeln ausgewertet, die Konzepte aufgrund der Benutzeraktionen bewerten, danach die Regeln, die Konzepte aufgrund der Bewertung ähnlicher Konzepte bewerten.

Die didaktische Komponente des Hilfesystems entscheidet danach, ob und wie das System intervenieren soll. Es gilt dabei abzuwägen: Zum einen soll das System möglichst unaufdringlich sein [J03], um den Lerner nicht vom eigentlich Inhalt abzulenken - andererseits sollte es dem Lerner bei groben Fehlern doch immer mit Hilfe zur Seite stehen. Deshalb besteht die

didaktische Komponente wiederum aus zwei Teilen: Neben der *intervention factory*, die basierend auf dem Overlay-Modell Interventionen vorschlägt, existiert noch das *intervention gate*, das unter Berücksichtigung der bisherigen Aktionen und der bereits ausgegeben Interventionen entscheidet, welche der Interventionen wichtig genug ist, um den Benutzer zu unterbrechen.

Wenn das Hilfesystem schließlich entscheidet, dass eine Intervention nötig ist, dann wird die letzte Aktion des Benutzers abgebrochen und der Hilfeagent erscheint und gibt den Inhalt der Intervention aus (Abb. 3). Da wir (optimistisch) davon ausgehen, dass der Lerner durch den erklärenden Text der Intervention tatsächlich etwas lernt, wird bei der Ausgabe der Intervention die Einstufung der Kenntnis der erklärten Konzepte im Benutzermodell verbessert.

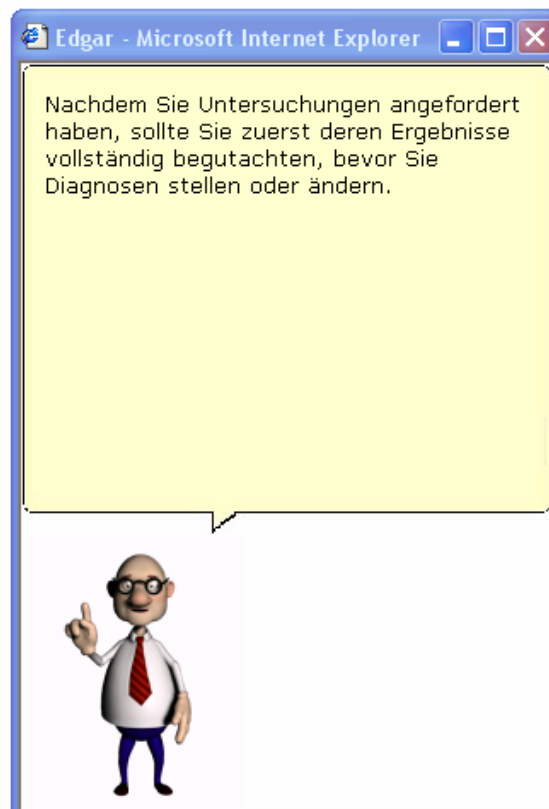


Abbildung 3: Der Hilfeagent

3 Evaluation

3.1 Evaluationsdesign

Im Rahmen des nach der neuen Approbationsordnung eingerichteten Kurses „Rheumatologie, Allergologie und Immunologie“ der Universität Würzburg wurde d3web.Train mit 30 Trainingsfällen (6 aus dem Gebiet Allergologie, 4 Immundefektsyndrome und 20 Rheumatologie) vorlesungsbegleitend eingesetzt. Der Kurs wurde von 62 Studierenden besucht, die überwiegend im 6. Semester waren. Die Studierenden mussten von den 30 Fällen mindestens 20 erfolgreich bearbeiten, um zur Klausur zugelassen zu werden. Da dies das erste Mal war, dass für die Klausur Übungsaufgaben bearbeitet werden mussten, durften die Studierenden alle Hilfen benutzen und den Fall auch mehrfach bearbeiten, wobei nur die beste Bewertung zählte. Die Fälle basierten dabei auf Arztbriefen [BBH04] und unterstützten nur die Aufgabenstellungen „Ergebnisse auswerten“ und „Diagnosen stellen“, da nach Einschätzung der Dozenten die Aufgaben „Untersuchungen anfordern“ und „Therapien festlegen“ noch zu anspruchsvoll für die Studierenden gewesen wären. Um das Hilfesystem zu testen, wurden die Studierenden hälftig in eine Studiengruppe mit eingeschaltetem, adaptivem Hilfesystem und eine Kontrollgruppen ohne adaptives Hilfesystem eingeteilt. Die Studierenden konnten nicht auf Anhieb erkennen, zu welcher Gruppe sie gehörten, da in der Benutzeroberfläche kein Hinweis auf das adaptive Hilfesystem enthalten ist und es sich nur sporadisch aus eigener Initiative meldet.

In beiden Gruppen wurden alle Bedienfehler protokolliert und darüber hinaus alle Aktionen des Benutzers, sowie Bearbeitungsdauer und Fallergebnis. Außerdem füllten die Studierenden zu Beginn einen kurzen Fragebogen aus und nach der abschließenden Klausur (deren Ergebnisse ebenfalls in die Evaluation einfließen) einen sehr ausführlichen. Für diese Evaluation werteten wir die Zufriedenheit mit dem Hilfesystem aus sowie die subjektiv eingeschätzte Einarbeitungszeit. Weitere Evaluationsergebnisse, insbesondere zur generellen Akzeptanz der Studierenden, finden sich in [RHTKBPK05].

3.2 Ergebnisse

Insgesamt wurden 2991 Fallbearbeitungen registriert, von denen 207 nicht abgeschlossen wurden. Unter den verbleibenden 2784 vollständigen Fallbearbeitungen gab es 1383 ohne Wiederholungen (maximal möglich wären $62 \cdot 30 = 1860$ gewesen; damit jeder Studierende genau die Klausurzulassung erhält $62 \cdot 20 = 1240$), die restlichen 1401 Fallbearbeitungen beruhten darauf, dass Studierende Fälle mehr als einmal bearbeiteten. Bei den Fragebögen bestand die Rücklaufquote für den kurzen Fragebogen zu Beginn 47 von 62 = 76%, beim abschließenden ausführlichen Fragebogen sogar 100%. Bei den 2991 Fallbearbeitungen wurden insgesamt 4091 Bedienfehler ($\approx 1,4/\text{Fall}$) automatisch registriert, wobei nur zwischen 3 Fehlerarten unterschieden wurde:

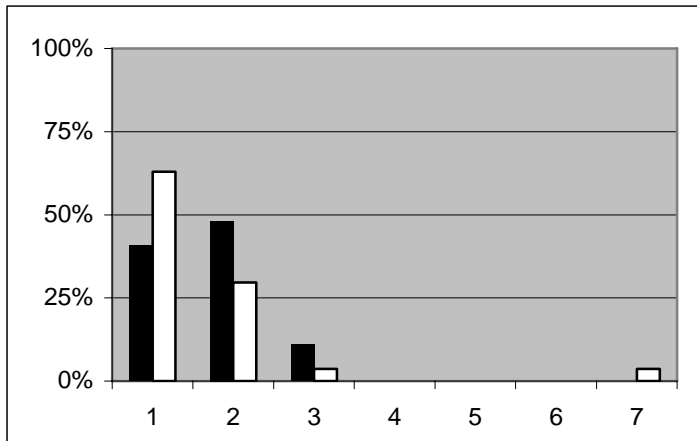
- Es wurden vorschnell Zwischen- oder Enddiagnosen eingegeben, ohne dass die bis dahin vorliegende Patientenakte vollständig angeschaut wurde (3433).
- In der Patientenakte vorhandene Bilder wurden nicht interpretiert (360).
- Es wurden vorschnell weitere Untersuchungen angefordert, bevor auf Basis der vorhandenen Untersuchungsergebnisse die erwartete und dem Benutzer angezeigte Anzahl und Art der Zwischendiagnosen gestellt wurde (298).

Im Folgenden präsentieren wir die Ergebnisse zur Effektivität und subjektiven Einschätzung des adaptiven Hilfesystems.

Einarbeitungszeit

Eine Frage des abschließenden Fragebogens war die nach der subjektiven Einarbeitungszeit. Die Studierenden sollten dabei angeben, nach welchem Zeitraum sie keine der nur anfänglichen Schwierigkeiten mehr hatten (manche Studierende hatten auch weiterhin Probleme mit dem System, dies wird aber noch genauer evaluiert). Die Studiengruppe mit eingeschaltetem Hilfesystem hatte signifikant früher den Eindruck, eingearbeitet zu sein (Tabelle 1).

Tabelle 1: „Wie lange haben Sie für die Einarbeitung in d3web.Train gebraucht?“ (1 = <10min, 2 = 10-30min, 3 = 30min-1h, 4 = 1-2h, 5 = 2-3h, 6 = 3-6h, 7 = >6h – weiss: Studiengruppe, schwarz: Kontrollgruppe; Rechts: Durchschnittswerte



Durchschnittswerte:

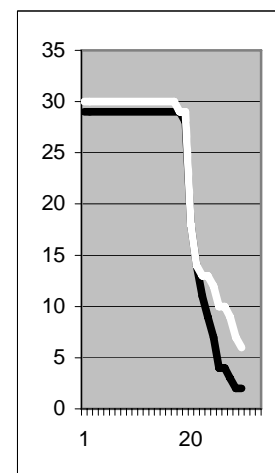
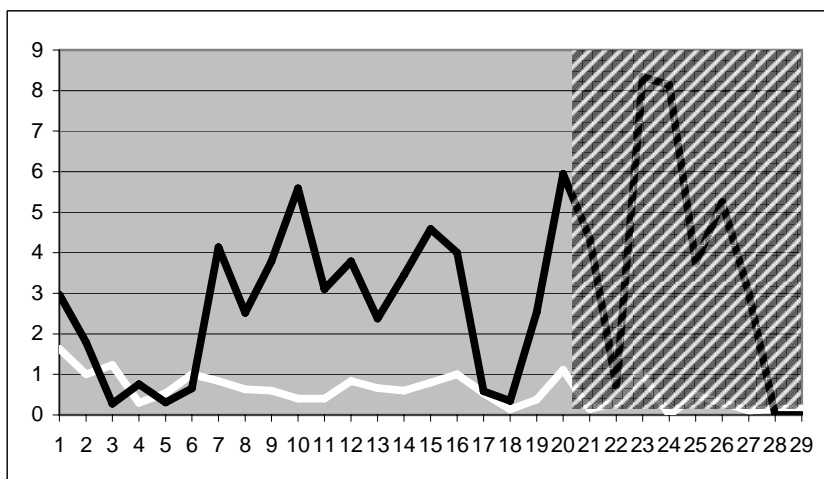
Studiengruppe:
1,6 (Standardabweichung:
1,2) entspricht ≈14min.

Kontrollgruppe:
2,0 (Standardabweichung:
1,0)
entspricht ≈20min.

Fehlerzahl

Tabelle 2 zeigt die Zuordnung der oben erwähnten Bedienungsfehler bezüglich der Studien- und Kontrollgruppe. Während die Studiengruppe im Schnitt nur 0,60 Fehler pro Fall machte, lag der Wert der Kontrollgruppe mit 3,00 5 mal so hoch. Zusätzlich zeigt die Tabelle den zeitlichen Verlauf der Fehlerquote an. Hier ist im Durchschnitt zwar ein anfängliches Absinken der Fehlerquote zu beobachten, die aber bei der Kontrollgruppe nach den ersten Fällen stark zunimmt. Die Schwankungen nach dem 20. Fall (schraffiert) liegen vermutlich daran, dass diese Fallbearbeitungen nicht mehr für die Klausurzulassung relevant waren und auch nur von wenigen Studierenden durchgeführt wurden.

Tabelle 2: Durchschnittliche Fehlerzahl bei der ersten Fallbearbeitung jeden Falles – weiss: Studiengruppe, schwarz: Kontrollgruppe – Insgesamt: Studiengruppe: 0,60 (Standardabweichung: 0,39), Kontrollgruppe: 3,00 (Standardabweichung: 2,30). Schraffierter Bereich: Fallbearbeitungen, die für die Klausurzulassung nicht mehr notwendig waren. Rechts: Teilnehmerzahl mit dieser Anzahl an Bearbeitungen.

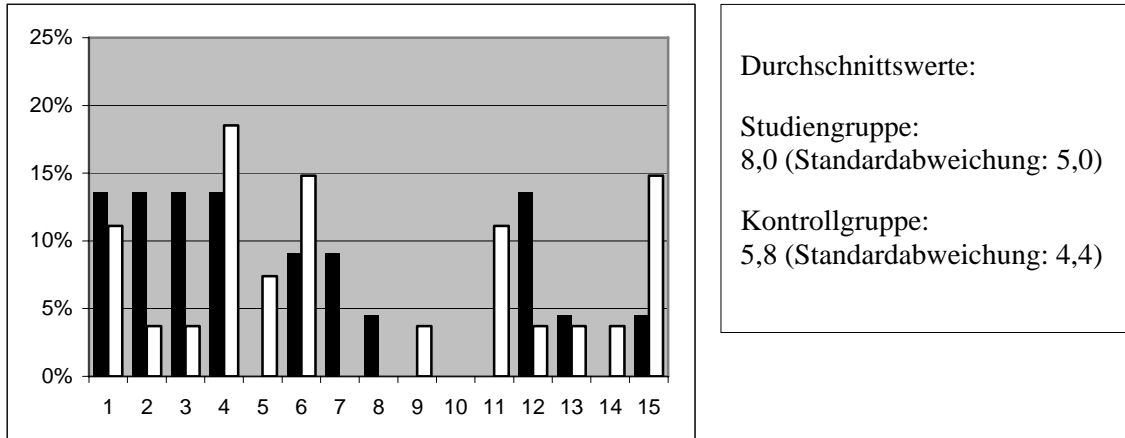


Bewertung des Hilfesystems

Eine weitere Frage des abschließenden Fragebogens bezog sich auf die subjektive Bewertung des Hilfesystems. Dazu ist anzumerken, dass auch die Studierenden der Kontrollgruppe ein Hilfesystem verfügbar hatten, dass sich jedoch nur auf inhaltliches Feedback bezog, während die Studiengruppe im Hilfesystem sowohl

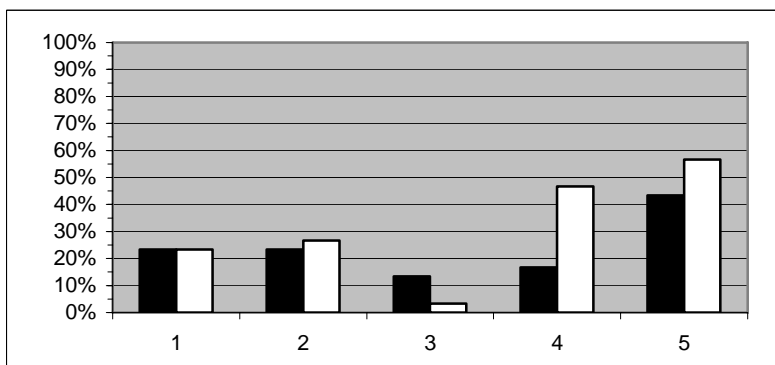
inhaltliches als auch auf die Bedienung bezogenes Feedback bekamen. Zusätzlich erschien bei beiden Gruppen in einer bestimmten Situation eine (hilfreiche) Meldung, die sich nur schwer von den Meldungen eines Hilfesystems unterscheiden ließ (allerdings wurde dabei nie eine Aktion des Lerner unterbrochen). Die Studiengruppe bewertete das Hilfesystem dabei genau in der Mitte zwischen „sinnvoll“ und „störend“ (8,0), während die Kontrollgruppe ihr (rudimentäres) Hilfesystem deutlich positiver bewerteten (5,8).

Tabelle 3: Antworten auf die Frage „Wie empfanden Sie die Meldungen des Hilfesystems?“ (1 = sinnvoll, ..., 15 = störend) – weiss: Studiengruppe, schwarz: Kontrollgruppe; Rechts: Durchschnittswerte



5 weitere Fragen im abschließenden Fragebogen bezogen sich auf verschiedene mögliche Varianten zur Konfigurierung des Hilfesystems und eine Aussage über das Hilfesystem:

Tabelle 4: Das Hilfesystem ... 1 = sollte seine Meldungen immer anzeigen, 2 = sollte die Meldungen einmal anzeigen dann lange nicht, 3 = sollte nie Meldungen anzeigen, 4 = hat Aktionen unpassend verhindert, 5 = sollte abschaltbar sein – weiss: Studiengruppe, schwarz: Kontrollgruppe.



3.3 Diskussion

Das adaptive Hilfesystem wurde entwickelt, um Benutzerfehler bei der Bedienung von d3web.Train zu verhindern. In Tabelle 2 zeigt sich, dass die Studiengruppe durch das Hilfesystem im Verlauf des Kurses sehr viel weniger Fehler gemacht hat als die Kontrollgruppe. Auch fühlten sich die Teilnehmer der Studiengruppe schneller eingearbeitet als die Kontrollgruppe (Tabelle 1), was wohl einerseits auf die erklärenden Texte der Interventionen zurückzuführen ist, andererseits merken die Benutzer eben auch durch das Ausbleiben von Interventionen, dass sie keine Fehler mehr machen.

In dieser Hinsicht war das Hilfesystem erfolgreich, auch wenn die Meinung der Teilnehmer dem teilweise widerspricht. Dies ist aber vermutlich größtenteils auf zwei Dinge zurückzuführen:

Einerseits verfolgten die Teilnehmer bei der Fallbearbeitung Ziele, die nicht vorhergesehen waren und die deshalb vom Hilfesystem eher behindert wurden, andererseits herrschte einige Verwirrung bzgl. der zu bewertenden Funktionalität.

Bei der wiederholten Bearbeitung eines Falles stand für die Teilnehmer nämlich meist nicht mehr der Fall als Ganzes im Vordergrund, sondern nur bestimmte Inhalte, z.B. Fragen zu bestimmten Bildern, das zu den Diagnosen verfügbare Hintergrundwissen mit den Diagnosekriterien oder der ausführliche Fallabschlusskommentar des Dozenten. Auf dem Weg zu diesen Informationen achteten die Teilnehmer aus der Kontrollgruppe (noch) weniger auf die korrekte Bearbeitung des Falles, die Teilnehmer aus der Studiengruppe mussten jedoch den Fall korrekt bearbeiten, da sonst das Hilfesystem intervenierte. Verständlicherweise wurde hier das Hilfesystem als störend empfunden, was sich besonders bei Antwort 4 („Das Hilfesystem hat Aktionen unpassend verhindert“) in Tabelle 4 niederschlägt (47%). Trotzdem waren noch 20% der Studiengruppe der Meinung, dass die Fehlermeldungen immer angezeigt werden sollten und nur 3% wollten das System ganz abgeschaltet haben. Die favorisierte Lösung war, dass das System vom Lerner abschaltbar sein sollte (57%). Eine diesbezüglich durchgeführte weitere Auswertung zeigte, dass die Fehlerzahl bei Teilnehmern der Studiengruppe bei der Wiederholung von bereits bearbeiteten Fällen deutlich höher lag (0,9 / Standardabweichung: 1,29) als bei der erstmaligen Fallbearbeitung, während die Fehlerrate bei den Studierenden der Kontrollgruppe sank (1,70 / Standardabweichung: 1,39) – auch ein Indiz dafür, dass der Fallabschlusskommentar für die Lerner wertvolle Informationen bot, denn dieser lässt sich einfach erreichen, indem man den Fall beendet (was jederzeit möglich ist und nicht als Fehler gezählt wird).

Allerdings ist nicht immer klar, welche Funktionalität die Teilnehmer jeweils bei ihren Aussagen bewerten. In der Diskussion zeigte sich, dass sowohl die Teilnehmer der Studiengruppe aber auch besonders die der Kontrollgruppe nicht zwischen den verschiedenen Meldungsarten differenzieren konnten, die der Hilfeagent anzeigt. Besonders die oben beschriebene „hilfreiche Meldung“ wurde überwiegend als störend empfunden (Ergebnis der Diskussion), diese konnte bei der Bewertung aber verständlicherweise nicht von den vom Hilfesystem generierten Meldungen unterschieden werden. Auch waren manche Teilnehmer aus der Kontrollgruppe enttäuscht, dass bei ihnen das Hilfesystem nicht aktiviert war und gaben als Antwort ihre Erwartung für ein solches System an. Möglicherweise stammen also die eher positiven Werte der Kontrollgruppe in Tabelle 3 von Teilnehmern, die das Hilfesystem gerne gehabt hätten und die eher negativen von solchen, die die „hilfreiche Meldung“ als besonders störend empfanden. Die Teilnehmer der Studiengruppe stuften das Hilfesystem insgesamt neutral ein, 15% empfanden die Fehlermeldungen als absolut störend (15) (obwohl sie bei durchschnittlich nur 0,6 Interventionen pro Fallbearbeitung eigentlich nur wenige Fehlermeldungen erhielten), 41% stuften die Fehlermeldungen aber als eher hilfreich ein (4-6), 19% sogar als hilfreich (1-3), 60% der Bewertungen waren also eher positiv.

Auch bei den Fragen zu verschiedenen Konfigurationen (Tabelle 4) zeigte sich wieder, dass auch Studierende aus der Kontrollgruppe eine Meinung dazu hatten – überraschenderweise hatten 17% der Kontrollgruppe den Eindruck, das (rudimentäre) Hilfesystem hätte bei ihnen Aktionen unpassend verhindert, obwohl dies technisch gar nicht möglich war. 13% dieser Gruppe wollten das Hilfesystem vielleicht deshalb deaktiviert haben vermutlich aber wegen der doch nicht so „hilfreichen Meldung“. Auch bei der Kontrollgruppe war ein großer Teil (43%) der Meinung, dass das Hilfesystem abschaltbar sein sollte.

4 Zusammenfassung und Ausblick

In diesem Papier haben wir vorgestellt, wie wir ein adaptives Hilfesystem einsetzen konnten, um die Bediener eines fallbasierten Trainingssystems beim Erlernen der Oberfläche zu unterstützen. Es zeigte sich, dass Lerner mit aktivierter Hilfe im Verlauf des Kurses weniger Fehler machten als diejenigen, bei denen die Hilfe nicht aktiviert war.

Es hat sich erst durch die Evaluation gezeigt, dass die Lerner das System auch intensiv zum Zugriff auf das Hintergrundwissen oder auf bestimmte Inhalte nutzen, eine Funktion, die noch nicht optimal unterstützt wird. Aus diesem Grund planen wir für den nächsten Kurs entsprechende Erweiterungen. Die Studierenden können dann einerseits bei bereits bearbeiteten Fällen schon vor dem Start des eigentlichen Falles auf das gesamte zugehörige Hintergrundwissen zugreifen, andererseits sollen sie den Fall auch mit komplett abgeschaltetem Hilfesystem durchspielen können.

Für den laufenden Kurs wurde bereits eine von den Studierenden (in der Diskussion) gewünschte Funktionalität implementiert: Der Lerner kann jede vom Hilfesystem generierte Meldung bewerten. Bewertet er sie schlecht, dann erscheint sie nicht wieder. Ehemals schlecht bewertete Meldung kann der Lerner aber auch

wieder aktivieren. Wir hoffen, dass die Lerner nicht voreilig Meldungen ausblenden und die Effektivität des Systems dadurch sinkt, andererseits dass dadurch die Akzeptanz des Hilfesystems steigt. Eine sinnvolle Erweiterung dieser Funktionalität könnte darin bestehen, dass das System selbständig eine Meldung reaktiviert, wenn der Teilnehmer immer wieder Fehler macht, die diese Meldung auslösen.

Literaturverzeichnis

- [BBH04] Betz, C., Buscher, H.-P., Hörnlein, A., Puppe, F., Schuhmann, M.: Generierung diagnostischer Trainingsfälle aus Arztbriefen, in [P04], 25-36, 2004.
- [CG77] Carr, B., Goldstein, I.P.: Overlays: A theory of modelling for computer aided instruction, Technical Report AI Memo 406, MIT, Cambridge, MA, 1977
- [d3web.Train] Fallbasierte intelligente Trainingssysteme für die Diagnostik, <http://www.d3webtrain.de> (01.08.2005).
- [FLS85] Fischer, G., Lemke, A., Schwab, S.: Knowledge-based Help Systems, Human Factors, in: Computing Systems – CHI'85 Conference Proceedings, pp. 161-167, ACM, New York, 1985
- [HBP02] Hörnlein, A.; Betz, C.; Puppe, F.: Redesign eines generativen, fallbasierten Trainingssystems für das WWW im d3web.Trainer, in: Bernauer, J.; Fischer, M. R.; Leven, J.; Puppe, F. and Weber, M. (eds.): Rechnergestützte Lehr- und Lernsysteme in der Medizin - Proceedings zum 6. Workshop der GMDS AG Computergestützte Lehr- und Lernsysteme in der Medizin, Fachhochschule Ulm, 11.-12. April 2002, ISBN 3-8322-0611-6, Shaker Verlag, Aachen, 2002
- [HP04] Hörnlein, A., Puppe, F.: Applying learner modelling for user interface assistance in simulative training systems, Proceedings of the workshop "Teaching and Learning Systems: The Role of AI" at the 27th German Conference on Artificial Intelligence (KI2004) in Ulm, 2004
- [HRKBP04] Hörnlein, A.; Reimer, S.; Kneitz, C.; Betz, C.; Puppe, F. Semantische Annotierung von Arztbriefen zur Generierung diagnostischer Trainingsfälle, in: accepted in Proc. Of Delfi-2004, "Lecture Notes in Informatics", 2004.
- [N01] Norman, D.A.: The Design of Everyday Things, ISBN 0-262-64037-6, MIT Press, 2001
- [P04] Pöpl, S., Bernauer, J., Fischer, M., Handels, H., Klar, R., Leven, J., Puppe, F., Spitzer, K. (eds.): Rechnergestützte Lehr- und Lernsysteme in der Medizin: Proc. 8. Workshop der GMDS AG Computergestützte Lehr- und Lernsysteme in der Medizin, Shaker, Aachen, 2004
- [R79] Reenskaug, T.: Models – Views – Controllers, Xerox PARC Technical Note, 1979
- [RHTKBP05] Reimer, S., Hörnlein, A., Tony, H.-P., Kraemer, D., Betz, C., Puppe, F., Kneitz, C.: Evaluation of a Case-Based Training System (d3web.Train) in Rheumatology, Poster CBT-Workshop gmds/dae 2005
- [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: [P04], pp. 155-164, 2004.
- [SHPM05] Schuhmann, M., Hörnlein, A., Puppe, F., Müller, J. G.: WOQ – Ein einfaches Autorensystem für skalierbare Bilderfragen-Quizsysteme im Web, in diesem Band
- [Skills Lab] SkillsLab – Medizin – Uni Würzburg, <http://www.skills-lab.medizin.uni-wuerzburg.de/> (01.08.2005).