Werkstattbericht

Semi-automatische Korrektur von juristischen Lösungsskizzen

Philip-Daniel Beck, Marianus Ifland, Jörn Kronbach, Frank Puppe, Ralf Schenke[*]

A. Einleitung

Im Rahmen des Jurastudiums stellt das Bearbeiten von juristischen Fällen ein wesentliches Lehr- und Lernziel dar. Um sich die hierfür notwendigen Fähigkeiten anzueignen, ist es unumgänglich, an möglichst vielen Sachverhalten zu üben. Dabei stellt die Korrektur der Bearbeitungen für die Studierenden ein wesentliches Mittel dar, um ihren Leistungsstand einschätzen zu können, die eigenen Fehler zu analysieren und dazuzulernen. Leider ist die manuelle Korrektur von Klausurlösungen zeit- und personalaufwendig. Dies führt im Jurastudium dazu, dass außerhalb von obligatorischen Leistungskontrollklausuren nur eine geringe Anzahl von Fällen zur Übung angeboten werden kann. Entsprechende Angebote beschränken sich zumeist auf den Examensklausurenkurs und damit die letzte Phase des Studiums. Um den Aufwand der Korrektur zu verringern, könnten statt kompletter Ausarbeitungen aber auch lediglich Lösungsskizzen betrachtet werden. Diese beschränken sich zwar auf die grobe Argumentationsstruktur und zentrale Stichworte. Gleichwohl kann aber auch schon anhand dieser gut erkannt werden, ob der richtige Lösungsweg eingeschlagen und die Problemschwerpunkte einer Klausuraufgabe herausgearbeitet wurden. Mit dem hier vorgestellten Tool, das an der Juristischen Fakultät der Universität Würzburg eingesetzt wird, wird eine semi-automatische Korrektur solcher Lösungsskizzen ermöglicht. Das Tool korrigiert die eingereichte Lösung durch einen Vergleich mit der Musterlösung zunächst automatisch. In einem zweiten Schritt kann diese Korrektur durch den menschlichen Korrektor nachbearbeitet werden. Durch eine abschließende, automatische Generierung eines Feedback-Dokumentes wird den Studierenden aufgezeigt, welche Teile ihrer Lösung gut oder verbesserungswürdig sind bzw. welche Teile der Lösung komplett fehlen. Finanziert wird das Projekt durch Studienzuschüsse, die in Bayern an die Stelle der früheren Studiengebühren getreten sind. Initiiert worden ist es im Zuge einer mittlerweile langjährigen Zusammenarbeit der Juristischen Fakultät mit dem Lehrstuhl für Künstliche Intelligenz und Angewandte Informatik, dem die Universität die Entwicklung und Implementierung zahlreicher E-Learning-Konzepte verdankt.

Teil B des Beitrages beschreibt mit dem entwickelten Tool verwandte Arbeiten, Teil C den Workflow und die allgemeine Funktionsweise des Tools. Anschließend wird eine Auswertung auf Grundlage von Daten aus den Sommersemestern in den Jahren 2014 und 2015 in Abschnitt D vorgestellt. Im letzten Abschnitt E fassen

* *Daniel Beck* und *Marianus Ifland* waren wissenschaftliche Mitarbeiter am Lehrstuhl von *Frank Puppe* für Künstliche Intelligenz und Angewandte Informatik der Universität Würzburg, *Jörn Kronbach* war wissenschaftlicher Mitarbeiter am Lehrstuhl von *Ralf Schenke* für Öffentliches Recht, Deutsches, Europäisches und Internationales Steuerrecht.

ZDRW 3/2016, DOI: 10.5771/2196-7261-2016-3-242

https://doi.org/10.5771/2196-7261-2016-3-242
Generiert durch IP '3.16.217.189', am 12.10.2024, 04:42:03.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

wir die Ergebnisse zusammen und diskutieren mögliche zukünftige Weiterentwicklungen.

B. Vorarbeiten und Stand der Entwicklung

Tools zur automatischen Korrektur von verschiedenartigen Aufgabentypen existieren viele. Dabei variieren die Korrekturaufgaben stark im Schwierigkeitsgrad, je nachdem, wie viele Freiheitsgrade der Lerner während der Lösung der gestellten Aufgabe bekommt. Relativ leicht ist die Korrektur für offene oder halboffene Aufgaben wie Single- oder Multiple-Choice- oder Lückentext-Aufgaben, da hier nur die Richtigkeit der gewählten Lösungen geprüft werden muss[1]. Wesentlich schwieriger ist die Korrektur von offenen Aufgaben, bei denen die Studierenden ihre Lösungen mit Hilfe natürlicher Sprache frei formulieren dürfen.

In der Literatur häufig untersuchte Aufgabentypen stellen Essays dar. Die Studierenden formulieren einen Text zu einer Fragestellung, welcher automatisch bewertet werden soll. Die Bewertung findet dabei anhand textueller Eigenschaften statt. Solche Eigenschaften sind zum Beispiel die Gesamtzahl an Wörtern, die Anzahl an Nebensätzen oder das Verhältnis von Groß- zu Kleinbuchstaben. Mit Hilfe dieser Features und manuell bewerteter Essays werden mathematische statistische Modelle gelernt, welche anschließend auf neue Essays angewendet werden. Problematisch dabei ist die häufige Vernachlässigung von inhaltlichen Aspekten des Essays. Beispiele für solche Bewertungstools sind *eRater*[2] oder *IntelliMetric*[3]. Dabei gibt es einige Studien, in denen Essays aus dem juristischen Bereich zur Evaluation herangezogen wurden.

Ein in Kooperation verschiedener Universitäten unter der Leitung der Fachhochschule Wismar entwickeltes Tool ist *T@keLaw*[4]. Hier werden sowohl Onlinevorlesungen angeboten, als auch Beispielfälle zur Verfügung gestellt, die gelöst werden sollen und anschließend automatisch korrigiert werden, wobei entsprechendes Feedback zur Verfügung gestellt wird. Die zu prüfenden Punkte sind dabei in einer Baumstruktur vorgegeben. Der Lerner[5] muss die entsprechend relevanten Punkte der Baumstruktur mit Entscheidungen versehen. Als Feedback bekommt der Lerner sowohl eine Punktzahl zurück, als auch ein Feedback zu einzelnen Prüfungspunkten. Systeme, die vorgefertigte Antworten zur Auswahl stellen, ermöglichen zwar eine Lernkontrolle, sind aber relativ weit von den Prüfungsanforderungen entfernt, mit denen die Studierenden im Staatsexamen konfrontiert sind. Hier be-

1 Ein entsprechendes System wird Studierenden der Universität Würzburg auf der Lernplattform CaseTrain angeboten (https://casetrain.uni-wuerzburg.de [22.07.2016]), die im Rahmen zahlreicher Vorlesungen und Übung der Juristischen Fakultät genutzt wird.

2 *Attali/Burstein*, in: ETS Research Report 2005, S. 1 (1 ff.).

3 *Rudern/Garcia/Welch*, in: GMAC Research Reports 2005, S. 1 (1 ff.).

4 *Müller*, in: Barton/Fuchs et al. (Hrsg.), S. 45 (45 ff.).

5 Soweit im Text das generische Maskulinum verwendet wird, erfolgt dies allein aus Gründen der besseren Lesbarkeit. Angesprochen und gemeint sind immer alle Menschen, gleich welchen biologischen oder sozialen Geschlechts.

https://doi.org/10.5771/2196-7261-2016-3-242
Generiert durch IP '3.16.217.189', am 12.10.2024, 04:42:03.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

steht die Schwierigkeit ja gerade darin, einen vollständigen Lösungsweg zu entwickeln, ohne dass der Studierende für die einzelnen Zwischenschritte ein Feedback erhält.

C. Beschreibung des Korrekturtools

Das hier vorgestellte Tool schlägt einen anderen Weg ein, indem die Studierenden eine vollständige Lösungsskizze erarbeiten müssen.

I. Aufgabe und Musterlösung

Die folgende, hier verkürzt wiedergegebene Beispielaufgabe wurde in den Sommersemestern 2014 und 2015 im Rahmen der Großen Übung für Öffentliches Recht gestellt:

> *In der bayerischen Stadt K erhebt die Lokalzeitung Korruptionsvorwürfe gegen die Stadtverwaltung. Im Stadtrat beantragt die G-Fraktion daraufhin die Einsetzung eines Sonderausschusses, um die Anschuldigungen aufzuklären. Gegen diesen Antrag, der mit 24 zu 23 Stimmen abgelehnt wird, stimmt auch die Stadträtin X, die bei dem Unternehmen, das von den Vorwürfen betroffen ist, als Angestellte beschäftigt ist.*
>
> *Stadträtin H erhebt vor dem Verwaltungsgericht Klage, um den Ablehnungsbeschluss aufzuheben oder hilfsweise seine Ungültigkeit festzustellen. Begründet wird die Klage mit der Befangenheit der X.*
>
> ***Aufgabe:*** *Hat die Klage Aussicht auf Erfolg?*

Einen Ausschnitt der Musterlösung, mit der die von den Studierenden erarbeiteten Lösungen abgeglichen werden, zeigt die Abbildung 1.

https://doi.org/10.5771/2196-7261-2016-3-242
Generiert durch IP '3.16.217.189', am 12.10.2024, 04:42:03.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

A. Sachurteilsvoraussetzungen / Sachentscheidungsvoraussetzungen / Zulässigkeit (+)[55]

I. Eröffnung des Verwaltungsrechtswegs, § 40 I 1 VwGO / § 40 Abs. 1 S. 1 / § 40 Abs. 1 Satz 1 VwGO / § 40 I S. 1 VwGO (+)[5]

1. aufdrängende Sonderzuweisung (-)(?)[0.5]

2. Generalklausel, § 40 I 1 / § 40 Abs. 1 S. 1 / § 40 Abs. 1 Satz 1 VwGO / § 40 I S. 1 VwGO (+)[4.5]

a) Öffentlich – rechtliche Streitigkeit (+)[1]

b) Nichtverfassungsrechtlicher Art (+)[2.5]

c) abdrängende Sonderzuweisung (-)[1]

II. Beteiligtenfähigkeit und Prozessfähigkeit / Beteiligten- und Prozessfähigkeit (+)*[20]**

1. Des Klägers H / Beteiligtenfähigkeit und Prozessfähigkeit des H / Beteiligten- und Prozessfähigkeit des H (+)

a) Beteiligtenfähigkeit des H / Beteiligtenfähigkeit § 61 VwGO(+)

b) Prozessfähigkeit des H / Prozessfähigkeit (+)

2. Des beklagten Stadtrats / Beteiligtenfähigkeit und Prozessfähigkeit des Stadtrats / Beteiligten- und Prozessfähigkeit des Stadtrats § 61 VwGO(+)

a) Beteiligtenfähigkeit des Stadtrats / Beteiligtenfähigkeit (+)

b) Prozessfähigkeit des Stadtrats / Prozessfähigkeit § 62 Abs. 3 VwGO (+)

III. Statthafte Klageart (+)*[30]**

1. Die statthafte Klageart richtet sich nach dem klägerischen Begehren gem. Art. 88 VwGO (+)

2. Anfechtungsklage, § 42 Abs. 1 VwGO / § 42 I VwGO (-)

3. Allgemeine Leistungsklage gerichtet auf Aufhebung des Beschlusses (-)

4. Klage Sui Generis (-)

5. Allgemeine Feststellungsklage, § 43 I VwGO / § 43 Abs. 1 VwGO (+)

6. Subsidiarität der allgemeinen Feststellungsklage, § 43 Abs. 2 VwGO / § 43 II VwGO (+)

Abbildung 1: Ausschnitt aus der zu Abbildung 1 zugehörigen Musterlösung (Erläuterung s. Text)

Die Musterlösung ist ebenso wie die eingereichten Lernerlösungen hierarchisch gegliedert. Als erster übergeordneter Prüfungspunkt muss die Zulässigkeit der Klage geprüft werden. Diese ist dann gegeben, wenn sämtliche Sachentscheidungsvoraussetzungen erfüllt sind, die auf der zweiten Gliederungsebene angesiedelt sind. Dazu gehören im Beispiel aus Abbildung 1 die „Eröffnung des Verwaltungsrechtswegs", die „Beteiligtenfähigkeit und Prozessfähigkeit" und die „Statthafte Klageart". Jeder dieser Punkte wird wiederum durch weitere Unterpunkte ausgefüllt, die hierarchisch der nächstniedrigeren Gliederungsebene zugeordnet sind. Jeder Gliederungspunkt wird mit einer Entscheidung in Form einer „(+)"- oder „(-)"-Auszeichnung

https://doi.org/10.5771/2196-7261-2016-3-242
Generiert durch IP '3.16.217.189', am 12.10.2024, 04:42:03.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

versehen. Diese gibt an, ob der entsprechende Punkt zutrifft oder nicht. Im Beispielsfall sind auf der ersten Gliederungsebene die Sachurteilsvoraussetzungen mit „(+)“ ausgezeichnet. Dies bedeutet, dass die Klage zulässig ist und jedenfalls die unmittelbar untergeordneten Gliederungspunkte ebenfalls vorliegen müssen und folgerichtig mit einem „(+)“ ausgezeichnet sind. In der Musterlösung sind die Gliederungspunkte mit Zahlenwerten in eckigen Klammern „[X]“ versehen. Mit diesen wird die Wichtigkeit des entsprechenden Punktes bewertet, so dass bei der Korrektur im Rahmen der Gesamtbewertung gewichtet werden kann. Um den Studierenden die Erstellung der Musterlösung zu erleichtern, steht auf der E-Learning Plattform ein Dokument zum Download bereit, in dem die einzelnen Gliederungsebenen als sich automatisch durchnummerierende Formatvorlagen voreingestellt sind[6]. Einmal eingegebener Text kann dann über einen Tastendruck sehr schnell der zutreffenden Gliederungsebene zugewiesen und gegebenenfalls noch korrigiert werden. Nach Fertigstellung einer Lösungsskizze kann diese über einen Upload in der zentralen E-Learning Plattform der Universität Würzburg abgegeben werden. Als Lösung akzeptiert wird sie nur, wenn sie formal korrekt, d.h. unter Beachtung der oben skizzierten Syntax erstellt worden ist.

II. Die Benutzeroberfläche für die Korrektur

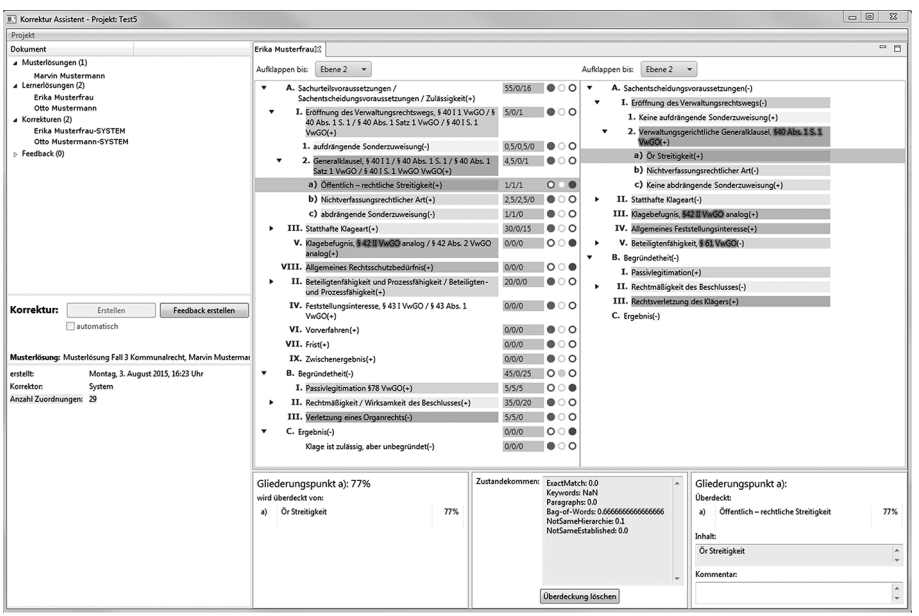

Abbildung 2: Benutzeroberfläche (GUI) (Namen verfremdet)

6 Das bereitgestellte Dokument ist in der Gliederungsansicht vom Textverarbeitungsprogramm Word formatiert und kann außer mit Microsoft Word auch mit dem freien Editor Open Office Writer bearbeitet werden, da beide das Zielformat .docx für die weitere Verarbeitung unterstützen.

https://doi.org/10.5771/2196-7261-2016-3-242
Generiert durch IP '3.16.217.189', am 12.10.2024, 04:42:03.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

Abbildung 2 zeigt die Benutzerfläche (GUI – graphical user interface) der erstellten Anwendung. Sie wurde in der objektorientierten Programmiersprache Java mit der frei verfügbaren Entwicklungsumgebung Eclipse e4-Framework[7] entwickelt. Unterteilt ist die GUI in verschiedene Bereiche (Views). Links oben befindet sich der Projektexplorer, in dem die verschiedenen eingereichten Lösungsskizzen der Studierenden in einer Baumstruktur verwaltet werden können. Darunter sind Schaltflächen zur Erstellung von Korrekturen und Feedbackdokumenten angeordnet. Links unten befindet sich ein Fenster für Meta-Informationen zu einer geöffneten Korrektur. Den Hauptteil der Anwendung nimmt der Editor zur Anzeige und Bearbeitung der Korrektur ein (mittlerer und rechter Bereich). Eine Korrektur wird durch Doppelklick auf ein Korrekturdokument im Projektexplorer gestartet. Im Editor werden Muster- und Lernerlösungsdokument parallel nebeneinander angezeigt. In der linken Hälfte des Editors befindet sich die Musterlösung, im rechten Teil die Lernerlösung. Beim Erstellen der Korrektur ordnet der Korrekturalgorithmus Segmente (also Unterpunkte) aus der Lernerlösung Segmenten aus der Musterlösung zu. Die Zuordnung wird farbig markiert, so dass gleichfarbig markierte Segmente auf beiden Seiten entsprechend zugeordnet sind. Zu jedem Segment der Musterlösung sind links daneben drei Zahlen angegeben. Diese geben die Wertigkeit des entsprechenden Segments an. Die erste Zahl steht für die maximale Wertigkeit des ganzen Teilbaums, also aller Segmente einschließlich des Untersegments. Die zweite Zahl gibt die Wertigkeit des Segments selbst an. Die dritte Zahl steht für die aktuelle Bewertung des Segments. Die Bewertung basiert auf der Qualität der Bewertung durch den Algorithmus. Diese ist durch die Farbe der Ampel rechts neben dem Segment angegeben. Ist die Ampel rot, so schätzt der Algorithmus diesen Punkt als schlecht bewertet ein, analog sind gelb und grün jeweils mittlere und gute Bearbeitungsqualität. Durch Klick auf die Ampel kann der Korrektor die Bewertungen der einzelnen Segmente ändern. Durch Markieren eines Segments lässt sich im View darunter sehen, mit welchem Segment auf der anderen Seite dieses assoziiert ist. Außerdem lässt sich bei Markierung zweier zugehöriger Segmente in der Mitte ablesen, wie diese zustande gekommen sind.

III. Ablauf der semi-automatischen Korrektur

Zur Vorbereitung der Korrektur muss der Korrektor im Tool ein neues Korrektur-Projekt erstellen, indem er Muster- und Lernerlösungen importiert. Der sich daran anschließende Korrekturvorgang gliedert sich in drei Schritte für jede Lernerlösung.

– Automatische Erstellung der Korrektur durch Vergleich einer Lernerlösung mit dem Musterlösungsdokument: Der Korrektor wählt die nächste zu korrigierende Lernerlösung aus. Hat er dies getan, erstellt er per Mausklick eine automatische Korrektur der Lernerlösung.

7 https://eclipse.org/e4/ (22.07.2016).

https://doi.org/10.5771/2196-7261-2016-3-242
Generiert durch IP '3.16.217.189', am 12.10.2024, 04:42:03.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

– Manuelle Überprüfung der erstellten Korrektur: Anschließend öffnet der Korrektor das erstellte Korrekturdokument und überprüft dieses manuell. Dazu geht er alle Punkte nacheinander durch und verbessert die automatische Korrektur seinen Wünschen entsprechend.
– Erstellung eines Feedbackdokumentes: Nach Abschluss der manuellen Überprüfung erstellt das Tool automatisch ein Feedbackdokument.

IV. Bewertungsschema

Bei der Korrektur soll vor allem die Struktur der abgegebenen Lernerlösungen bewertet werden. Dadurch wird vermieden, dass Studierende einzelne Punkte im Detail ausarbeiten müssen. Entscheidend ist, dass die relevanten Prüfungspunkte in der passenden Struktur abgearbeitet werden. So ist es zum Beispiel zur Überprüfung des Punktes „Eröffnung des Verwaltungsrechtswegs“ notwendig, die entsprechenden Unterpunkte „Abdrängende Sonderzuweisungen“ sowie „die Generalklausel § 40 Abs. 1 Satz 1 VwGO“ mit entsprechenden Unterpunkten zu prüfen. Dabei können einzelne Unterpunkte stärker in die Wertung eingehen als andere. Die Wertigkeit gibt der Korrektor in der Musterlösung an. Außer dem Vorhandensein bestimmter Prüfungspunkte in der geforderten Struktur wird die Entscheidung eines Punktes (d.h. die Auszeichnung mit (+) bzw. mit (-)) für dessen Bewertung herangezogen. Konnte ein Punkt zwar zugeordnet werden, aber ist die Entscheidung falsch, so ist der Punkt auch falsch bearbeitet.

V. Automatische Erstellung der Korrektur

Die automatische Erstellung der ersten Korrektur geschieht auf Grundlage eines Vergleichs zwischen Muster- und Lernerlösung. Dabei findet ein hierarchisches Matching statt, das versucht, das obige Bewertungsschema abzubilden. Auf der ersten Hierarchieebene werden Segmente verschiedenen Kategorien zugeordnet. Diese werden durch die Musterlösung vorgegeben. So gibt es im Fall Verwaltungsrecht die Kategorien „Sachentscheidungsvoraussetzungen“, „Begründetheit“ und „Ergebnis“. Damit Segmente aus der Lernerlösung zugeordnet werden können, müssen diese der gleichen Kategorie angehören. Dies ist nur dann der Fall, wenn diese ebenfalls unter dem Namen der Kategorie im Lernerlösungsdokument angeordnet sind. Über Kategoriengrenzen hinweg findet kein Matching statt. Im nächsten Schritt werden Teilbäume der zweiten Ebene (z.B. „Eröffnung des Verwaltungsrechtswegs“, „Statthafte Klageart“ usw.) zugeordnet. Dabei wird für jedes Segment dieser Ebene das am besten passende Segment aus der Musterlösung gesucht. Um dieses zu finden, vergleicht der Algorithmus einen Baum auf zweiter Ebene mit jedem Baum gleicher Kategorie und selber Ebene aus dem Musterlösungsdokument. Um einen Matching-Wert zu berechnen, werden alle Segmente aus den Teilbäumen paarweise gegeneinander verglichen. Für jedes Segment geht der beste Match in die durchschnittliche Gesamtbewertung dieses Teilbaums ein. Der Teilbaum mit dem höchsten Matching-Wert wird anschließend zugewiesen. In

https://doi.org/10.5771/2196-7261-2016-3-242
Generiert durch IP '3.16.217.189', am 12.10.2024, 04:42:03.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

Abbildung 2 ist ein Beispiel zu sehen, wie die einzelnen Punkte zugeordnet werden konnten. Die Bewertungen für die einzelnen Punkte der Musterlösung werden dann entsprechend von unten nach oben berechnet. Konnte ein Punkt aus der Musterlösung einem Punkt aus der Lernerlösung zugeordnet werden und ist die Entscheidung als richtig erkannt worden, so wird der Punkt mit „Grün“ bewertet. Andernfalls erscheint der Punkt „Rot“. Bei hierarchisch übergeordneten Punkten wird ein gewichteter Durchschnitt der Unterpunkte berechnet und „Grün“, „Gelb“ oder „Rot“ eingefärbt.

Das Matching zwischen Segmenten erfolgt anhand verschiedener Aspekte. Dabei wird jeweils für Paare von Segmenten ein Score zwischen 0 und 1 berechnet. Je näher der Wert bei 1 liegt, desto höher ist die angenommene Ähnlichkeit. Eine besondere Stellung für das Matching nehmen Paragraphen wie z.B. „§ 40 Abs. 1 Satz 1 VwGO“ ein. Wird in einem Paar von Segmenten derselbe Paragraph genannt, so wird der Score auf 1 gesetzt. Beim Matching wird auch darauf geachtet, dass Paragraphen unterschiedliche Schreibweisen haben können, so dass auch die Schreibweisen „§ 40 Abs. 1 Satz 1 VwGO“ und „§ 40 I S. 1 VwGO“ als identisch erkannt werden. Wenn kein gleicher Paragraph gefunden wurde, werden die normalisierten, bedeutungstragenden Wörter der beiden Segmente verglichen, indem berechnet wird, wie viel Prozent der Wörter der Lernerlösung in der Musterlösung vorkommen und umgekehrt. Dabei können auch Synonyme berücksichtigt werden. Zusätzlich wird die getroffene Entscheidung (+) oder (-) und die Stellung in der Hierarchieebene in den Score mit einbezogen.

VI. Feedbackdokumenterstellung

Nach Abschluss der manuellen Verbesserung der Korrektur wird auf dieser Grundlage ein Feedbackdokument erstellt. Dabei wird versucht, die Korrektur möglichst nochmals zusammenzufassen sowie die korrigierten Elemente entsprechend genau abzubilden. Als Format des Feedbackdokumentes wurde ebenfalls Word gewählt, da bisher kein Web-Interface vorhanden ist und Word Dokumente mit Hilfe der Software-Bibliothek docx4j relativ gut automatisch erstellt werden können. Ein Ausschnitt aus einem Feedbackdokument ist in Abbildung 3 dargestellt. Es zeigt die Lösung eines Studierenden zusammen mit der entsprechenden Musterlösung. Die zugeordneten Elemente werden entsprechend farbig markiert, so dass er seine Lösung leicht in Relation zur Musterlösung setzen kann. Außerdem werden die Qualitätsbewertungen eines jeden Segments in entsprechende Statuskommentare übersetzt, also „Rot“ nach „Fehler!“ bei gefundener Zuordnung oder nach „Fehlt!“, falls keine Zuordnung gefunden werden konnte, „Gelb“ nach „Geht noch besser!“ und „Grün“ nach „Gut!“. Zu jedem Segment der Lernerlösung kann der Korrektor manuell Kommentare eintragen. Sind solche vorhanden, werden diese in der entsprechenden Spalte angezeigt. Das Feedback enthält die dargestellte Zuordnung einmal aus Sicht der Musterlösung und einmal aus Sicht der Lernerlösung, so dass beide Fälle vom Studierenden leicht nachvollzogen werden

https://doi.org/10.5771/2196-7261-2016-3-242
Generiert durch IP '3.16.217.189', am 12.10.2024, 04:42:03.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

können. Weiterhin enthält das Feedbackdokument eine Zusammenfassung der Korrektur, in der eine Gesamtbewertung in Form einer Wertung zwischen 0 und 100 gegeben wird. Dies geschieht auf Grundlage der einzelnen Bewertungen. Zusätzlich wird für besonders relevante Punkte ausgegeben, ob diese gut oder schlecht gelöst wurden.

DETAILLIERTE BEGRÜNDUNG AUSGEHEND VON IHRER LÖSUNG:

SACHENTSCHEIDUNGSVORAUSSETZUNGEN

Deine Lösung	Kommentar	Musterlösung	Status
I. Eröffnung des Verwaltungsrechtswegs (-)		I. Eröffnung des Verwaltungsrechtswegs, § 40 I 1 VwGO / § 40 Abs. 1 S. 1 / § 40 Abs. 1 Satz 1 VwGO / § 40 I S. 1 VwGO (+)	Fehler!
1. Keine aufdrängende Sonderzuweisung (+)		1. aufdrängende Sonderzuweisung (-)	Fehler!
2. Verwaltungsgerichtliche Generalklausel, §40 Abs. 1 S. 1 VwGO (+)		2. Generalklausel, § 40 I 1 / § 40 Abs. 1 S. 1 / § 40 Abs. 1 Satz 1 VwGO / § 40 I S. 1 VwGO VwGO (+)	Fehler!
a) Ör Streitigkeit (+)		a) Öffentlich – rechtliche Streitigkeit (+)	Gut!
b) Nichtverfassungsrechtlicher Art (-)		b) Nichtverfassungsrechtlicher Art (+)	Fehler!
c) Keine abdrängende Sonderzuweisung (+)		c) abdrängende Sonderzuweisung (-)	Fehler!

Abbildung 3: Ausschnitt aus Feedbackdokument

D. Evaluation

In einer ersten Evaluation wurden 34 Lernerlösungen aus dem Sommersemester 2015 evaluiert. Im Schnitt enthielt eine Lernerlösung 27 Segmente, wobei der Umfang der Lösungen zwischen 16 und 41 Segmenten schwankt. Die Musterlösung bestand aus 39 Segmenten. Wie oben beschrieben, wurden die Dokumente zuerst automatisch durch das Tool vorkorrigiert. Anschließend wurden die Korrekturen durch einen juristisch qualifizierten Korrekturassistenten manuell nachkorrigiert. Dabei wurden die vorgenommenen Veränderungen automatisch mitgezählt. Einerseits betrifft dies Veränderungen der Bewertung eines Segmentes, andererseits Änderungen der Zuordnungen.

In zwei Fällen war eine sinnvolle Korrektur überhaupt nicht möglich. Auf Grund stark abweichender schematischer Strukturen der Lernerlösungen konnte das Tool in diesen beiden Fällen keine sinnvolle Vorkorrektur erstellen. Insofern beziehen sich die folgenden Zahlen nur auf die 32 korrigierbaren Fälle.

Verändert wurde die Bewertung der Korrektur zu 56 aus 1248 Segmenten. Damit lässt sich sagen, dass der Algorithmus in 95,5 % der Fälle mit der automatischen Korrektur richtig lag. Die Anzahl an vorgenommenen Veränderungen ist unterschiedlich. So wurden maximal 8 Veränderungen der Bewertung in einem Dokument vorgenommen. In 4 Dokumenten wurde hingegen gar keine Veränderung der Bewertung vorgenommen, diese können also als vollständig automatisch korrigiert angesehen werden.

Insgesamt wurden 838 Zuordnungen von Segmenten in Lernerlösungen zu Segmenten in Musterlösungen automatisch gefunden, die vom Korrektor akzeptiert

https://doi.org/10.5771/2196-7261-2016-3-242
Generiert durch IP '3.16.217.189', am 12.10.2024, 04:42:03.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

wurden, wobei in den oben erwähnten 56 Segmenten die Bewertung der Zuordnung verändert wurde.

Die Korrekturzeit unter Verwendung des Tools betrug pro Korrektur im Schnitt circa 7 Minuten.

E. Ausblick

Die bisherige Arbeit und Evaluation haben einen ersten Nachweis für die Nützlichkeit des entwickelten Tools geliefert. In zukünftigen Untersuchungen soll dessen Funktionsumfang allerdings noch um eine ganze Reihe weiterer nützlicher Eigenschaften erweitert werden. Im ersten Schritt soll der Vergleich mit mehreren Musterlösungen eingeführt werden. Dies ist notwendig, da juristische Lösungsskizzen nicht eindeutig sind und durch Vergleiche mit nur einer Musterlösung keine optimale automatische Korrektur für alle Lösungen erreicht werden kann. Weiterhin sollen zukünftig Segmente aus Lernerlösungen bei entsprechender Qualität als Mustersegmente definiert werden können. Dadurch soll eine dynamische Erstellung der Musterlösung aus den Musterlösungsdokumenten und den entsprechenden Lernersegmenten möglich sein. Gelingt dies, könnte für jedes Segment das beste Musterlösungssegment aus einer Menge vorhandener Mustersegmente erstellt werden. Aufbauend darauf sollen auch Lernersegmente, die sich nicht als Mustersegment eignen, für die weitere Korrektur herangezogen werden. Im Idealfall können Korrekturen aus bereits korrigierten Segmenten für neue zu korrigierende Segmente hergeleitet werden.

Ursprünglich war es geplant, zusätzlich zur Bewertung der Gliederung auch genauere Argumentationen und Begründungen, in denen Lerner ihre Entscheidungen weiter ausführen und rechtfertigen, zu bewerten. Da dies die Komplexität der Problemstellung weiter erhöht hätte, bleibt dies ebenfalls zukünftigen Fortentwicklungen des Korrekturtools vorbehalten. Endziel des Projekts ist eine Bereitstellung für die Öffentlichkeit. Wer das Tool gleichwohl bereits im Forschungsstadium ausprobieren will, kann sich gerne mit dem Autorenteam in Verbindung setzen[8].

Literaturverzeichnis

Attali, Yigal/Burstein, Jill, Automated essay scoring with e-rater V.2.0, in: ETS Research Report 2005, Reportnummer RR-04-45.

Müller, Tony, Die Rechtsvorlesung in WI: Neue Möglichkeiten durch E-Learning, in: Barton/Fuchs/Kuhn/Lämmel/Müller (Hrsg.), E-Learning-Inhalte: effizient entwickeln und effektiv einsetzen 2009, S. 45-82.

Rudern, Lawrence M./Garcia, Veronica/Welch, Catherine, An Evaluation of IntelliMetric™ Essay Scoring System Using Responses to GMAT® AWA Prompts, in: GMAC Research Reports (2005), Reportnummer RR-05-08.

8 Wir bitten, entsprechende Anfragen per E-Mail an frank.puppe@uni-wuerzburg.de bzw. schenke@jura.uni-wuerzburg.de zu richten.

https://doi.org/10.5771/2196-7261-2016-3-242
Generiert durch IP '3.16.217.189', am 12.10.2024, 04:42:03.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.