

Automatische Bewertung von Datenbankaufgaben unter Verwendung von LON- CAPA und Praktomat

Marcel Kruse, Nils Jensen
Zentrum für erfolgreiches Lehren und Lernen
Ostfalia Hochschule für angewandte Wissenschaften
{m.kruse, n.jensen}@ostfalia.de

Abstract

In dieser Ausarbeitung wird eine, im Rahmen des eCult-Projektes erarbeitete, Lösung zur automatischen Bewertung von Datenbankaufgaben mittels des LCMS LON-CAPA sowie des externen Bewertungstools Praktomat vorgestellt. Hierbei soll zum Einem auf die Verbindung des LCMS mit externen Bewertungstools, wie dem Praktomat, sowie auf die Funktionsweise und das Verfahren der automatischen Bewertung von SQL-Aufgaben näher eingegangen werden.

1 Einleitung

MINT-Lehrende sind jedes Semester erneut mit einer überwältigenden Zahl an begleitenden Übungs- und Labor-Aufgaben konfrontiert. Das Paper spezifiziert ein erprobtes System zum Bewerten von Einreichungen in SQL, das begleitende Aufgaben online zur Verfügung stellt. Es speichert diesbezügliche Einreichungen und bewertet sie automatisch. Es verwendet ein Oracle DBMS zur gesicherten Evaluation der SQL-Kommandos und wertet dessen Ausführungsergebnisse aus, um zur Bewertung zu gelangen.

Unser System ist neu, weil es existierende Lernräume mit automatischen Bewertern verbindet und so deren Vorteile bei der Bereitstellung von SQL-Aufgaben vereint:

- Im Lernraum sind die studentischen Stammdaten und die Erfolgsstatistiken zur Evaluation durch Lehrende und Studierende abgelegt. Weiterhin finden sich dort die in Kursen organisierten Online-Aufgaben. Die Aufgaben können zwischen den Lehrenden wiederverwendet werden.
- Automatische Bewerter nehmen die Rolle von Korrektoren an, die die Syntax und Semantik eines Computerprogramms benoten, z. B. in Java, Python oder SQL.

Automatische Bewerter wie z. B. der Praktomat [1] parsen den Quelltext eines Programms und führen Black-Box-Testfälle auf dessen gestarteten Kompilat aus. Dadurch ergibt sich die Möglichkeit, White- und Black-Box-Tests auszuführen. Hierzu werden vom Praktomat wiederum eine Reihe von existierenden Plugins zur Softwarequalitätssicherung verwendet.

Eine Herausforderung, deren Bewältigung uns durch den Praktomat abgenommen wird, liegt darin, dass das zu bewertende Programm unsicher ist und daher mittels Timeouts und Sandboxing abgesichert werden muss. Das dient dazu, z. B. Endlosschleifen, das Versenden von Emails oder das Kapern des Servers zu unterbinden.

Im engeren vergleichbare Systeme zu dem von uns vorgestellten System haben VIOPE Ltd.¹ und unter anderem die Hochschule Hannover erarbeitet [2]. Jedoch erlaubt derzeit

¹viope.com/en/schools/products/viope-courses.html

nur unsere Lösung den weltweiten elektronischen Austausch der SQL-Aufgaben in dem Peer-to-Peer-Netzwerk loncapa.org.

Kapitel 2 spezifiziert die verwendeten Systeme. Kapitel 3 beschreibt den technischen Ablauf. Kapitel 4 gibt den Ausblick.

2 LON-CAPA und Praktomat

In den folgenden beiden Kapiteln soll die Möglichkeit der eingangs erwähnten automatischen Bewertung von Datenbankaufgaben unter Verwendung des Learning Content Management Systems (LCMS) LON-CAPA und des Bewertungstools Praktomat diskutiert und erläutert werden.

Mit dem LCMS LON-CAPA existiert bereits ein System zur automatischen Bewertung von, unter anderem, mathematischen Aufgaben unterschiedlichster Ausprägung. Aufgrund des Open-Source Charakters von LON-CAPA werden darüber hinaus ständig Weiterentwicklungen erarbeitet, um das Spektrum der automatisch bewertbaren Aufgaben stetig zu erweitern. Hierbei werden unter anderem externe Grading-Tools angebunden, die solche Bewertungen, beispielsweise, auch für von LON-CAPA nicht nativ unterstützte Aufgaben zu Programmiersprachen ermöglichen sollen. Exemplarisch lassen sich hierfür Aufgaben in den Programmiersprachen Java und Python nennen. Ein Tool, was die Bewertung derartiger Aufgaben ermöglicht, stellt beispielsweise das System Praktomat dar, das mittels einer REST-Schnittstelle mit LON-CAPA verbunden ist. Hierbei werden die studentischen Einreichungen zur Lösung einer etwaigen Aufgabe, mittels eines Post-Requests an den Praktomat gesendet. Die Antwort des externen Graders erwartet LON-CAPA in XML-Format. Das Sequenzdiagramm in Abbildung 1 veranschaulicht diesen Austausch zwischen LON-CAPA und Praktomat nochmals etwas genauer.

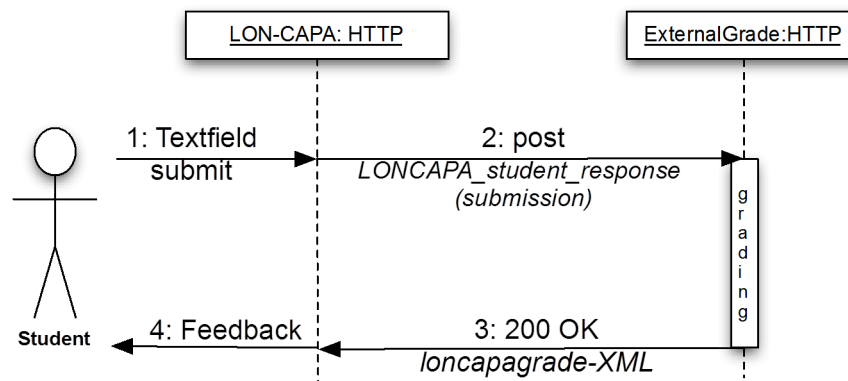


Abbildung 1: Sequenzdiagramm der externen Bewertung

Der Praktomat nimmt eine automatische Bewertung von studentischen Lösungen für Programmieraufgaben aus einem LON-CAPA-Kurs vor, in dem die Studierenden eingeschrieben sind. Durch Hochladen und Einreichen ihrer Lösung in LON-CAPA stoßen die Studierenden den Prozess auf Seiten des Praktomat an, der die Bewertung der eingereichten Lösung veranlasst. Das Ergebnis dieser Bewertung wird anschließend an LON-CAPA zurück gegeben. All diese Prozesse laufen vollständig im Hintergrund ab, so dass die Studierenden nichts davon erfahren. Aus Benutzersicht ist diese Bewertung somit identisch mit der Bewertung einer solchen Aufgabe, die vollständig auf Seite von LON-CAPA vorgenommen wird. Die vom Praktomat zurückgelieferte Auswertung stellt hierbei eine vorläufige Bewertung dar. Die endgültige Bewertung kann von Lehrenden im Kurs noch nachträglich angepasst und verändert werden, um etwaige Szenarien, die eine abweichende Bewertung nötig machen, abdecken zu können.

Die Struktur und das Verhältnis von LON-CAPA und dem Praktomat zueinander werden in Abbildung 2 genauer dargestellt.

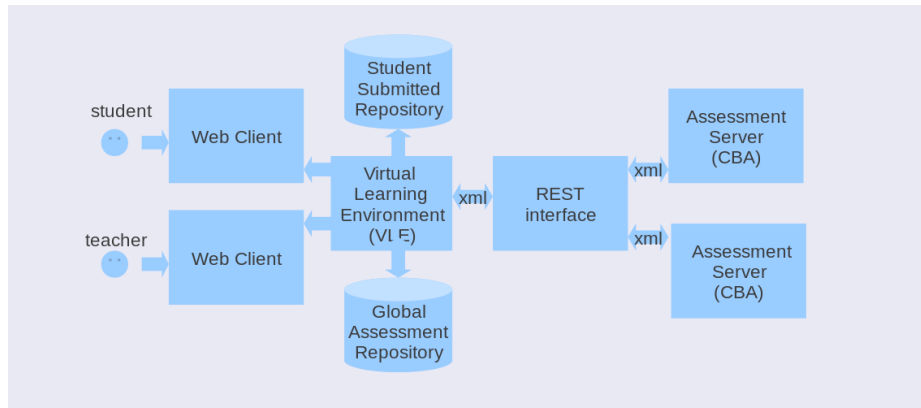


Abbildung 2: Schema der Praktomat-Anbindung

3 Bewertungsverfahren für SQL-Aufgaben

Über die Bewertung von Java-Aufgaben hinausgehend ist es, mit Hilfe des Praktomat, nun auch möglich, SQL-Aufgaben automatisch bewerten zu lassen. Um dies zu erreichen, wurde für die Verwendung auf dem Praktomat, in der Skriptsprache Python ein Script-Checker entwickelt, dessen Funktionsweise im Folgenden näher erläutert wird. Hierbei soll zum Einen die Struktur des zugrunde liegenden SQL-Interpreters erklärt werden. Darüber hinaus soll zum Anderen auf den Ablauf und das Verfahren der Auswertung an sich eingegangen werden. Zur besseren Illustration erfolgt dies auf Grundlage des Flussdiagramms in Abbildung 3, welches einen Überblick über das Verfahren der eigentlichen Bewertung der studentischen Lösung gibt.

Der Checker für die Bewertung der SQL-Aufgaben nimmt zunächst eine, auf dem Praktomat zur Verfügung stehende, Musterlösung einer, in LON-CAPA existierenden, SQL-Aufgabe entgegen.

(1) Diese wird dahingehend verarbeitet, dass sie zunächst in einzelne SQL-Befehle zerlegt wird.

(2) Die von den Studierenden über LON-CAPA eingereichte Lösung der Aufgabe wird anschließend auf identische Weise zerlegt.

(3) Daraufhin wird jeder der SQL-Befehle aus der Musterlösung separat an einen SQL-Interpreter weitergeleitet, der die Einreichung auf Syntax-Fehler untersucht. Das Ergebnis dieser Prüfung wird mittels des Skript-Checkers verwaltet.

(4) Gleiches geschieht analog für jeden einzelnen Befehl der studentischen Einreichung.

Der in diesem Zusammenhang verwendete Interpreter, ist ein Oracle DBMS mit einem von Prof. Dr. J. S. Lie et al entwickelten Web-Interface. Er erhält die zu bewertenden Befehle mittels eines get-Requests über das Web-Interface und generiert, nach Ausführung des betreffenden Befehls auf einer im Hintergrund laufenden Datenbank, eine Antwort-Seite, die das Ergebnis der Ausführung des Befehls enthält.

Da alle Datenbank-Transaktionen des Web-Interpreters mit derselben Datenbank ablaufen, muss vom Script sichergestellt werden, dass nach der Bewertung der Musterlösungen sämtliche erzeugten Tabellen wieder gelöscht werden. Dies hat den Hintergrund,

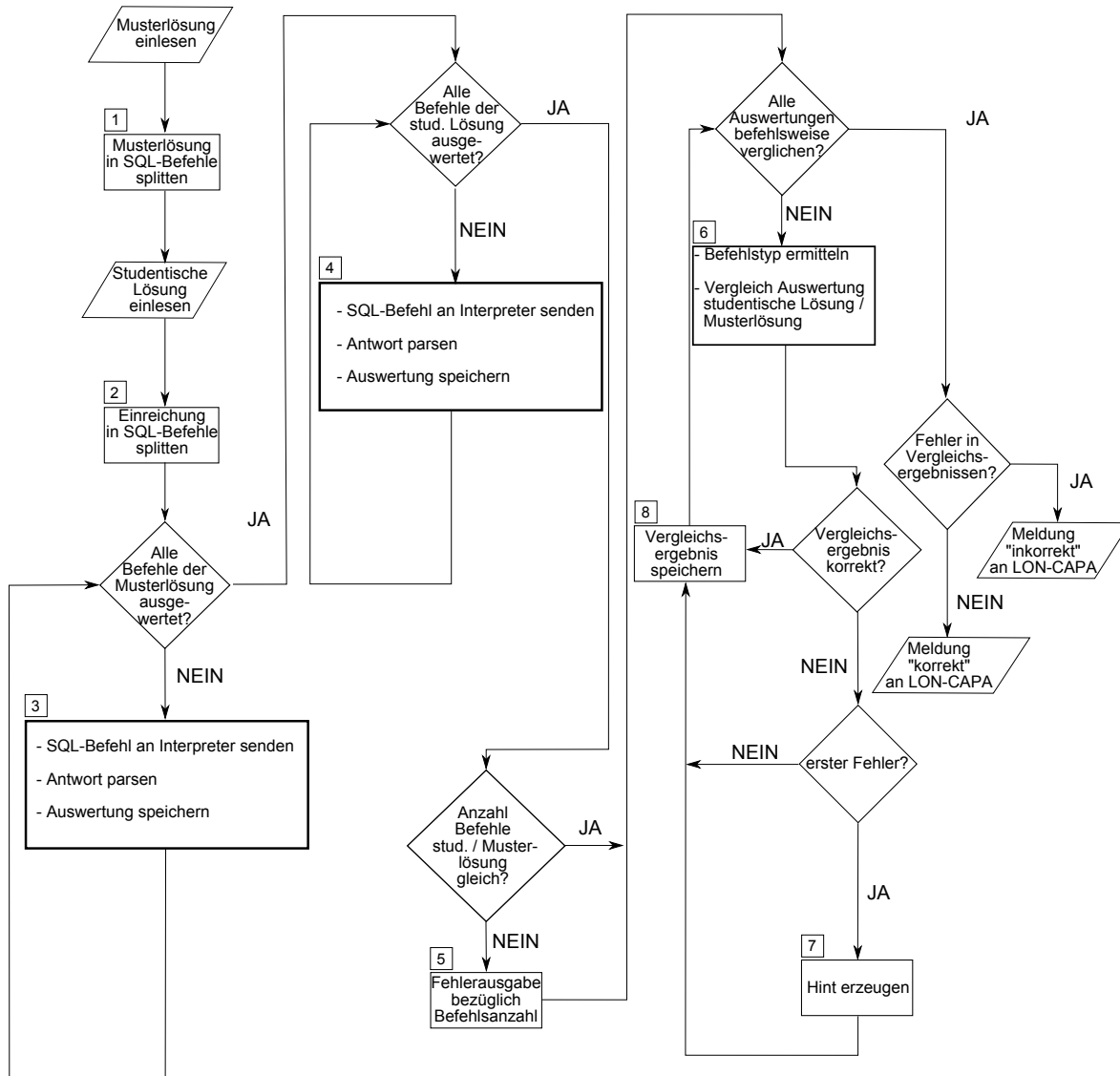


Abbildung 3: Bewertungsverfahren

falsch negative Ergebnisse, wie fehlgeschlagene "CREATE"-Anweisungen aufgrund bereits existierender Tabellen, zu verhindern. Analog dazu wird darüber hinaus, vor dem Beginn der automatischen Bewertung, geprüft, ob eine so genannte "Lock-Table" in der Datenbank vorhanden ist, um zu verhindern, dass mehrere Bewertungsvorgänge gleichzeitig auf das Web-Interface zugreifen. Existiert die Tabelle bereits in der Datenbank, versucht das Skript mehrfach diese "Lock-Table" anzulegen, und die Kommunikation mit dem Interface für den aktuellen Bewertungsvorgang zu reservieren. Nach einer festgelegten Anzahl von fehlgeschlagenen Versuchen erfolgt ein Time-Out mit der Rückmeldung an LON-CAPA, dass die Einreichung zu einem späteren Zeitpunkt wiederholt werden soll. Nach Abschluss der automatischen Bewertung, wenn alle Befehle der studentischen Einreichung und der Musterlösung an das Interface gesendet und ihre Auswertung gespeichert wurden, wird auch die "Lock-Table" gelöscht, um das Interface für den nächsten Bewertungsvorgang frei zu geben.

In einem nächsten Schritt wird überprüft, ob die Anzahl der eingereichten Befehle aus Musterlösung und studentischer Einreichung an den Interpreter übereinstimmt.

(5) Ist dies nicht der Fall, erfolgt eine Fehlerausgabe mit einem Hinweis auf die Differenz der Befehlsanzahl.

(6) Im Anschluss geschieht, skriptseitig, ein befehlsweiser Vergleich der Feedbacks des Interpreters. Hierbei wird eine Unterscheidung zwischen Feedbacks zu verschiedenen Typen der jeweiligen SQL-Statements, sowie eine diesem entsprechende Auswertung vorgenommen. Dies stellt dahingehend eine Überprüfung der studentischen Einreichung dar, dass sichergestellt wird, ob diese nicht nur syntaktisch korrekt ist, sondern gleichsam auch inhaltlich eine valide Lösung der Aufgabenstellung widerspiegelt. Für den Fall, dass bereits der SQL-Interpreter einen Syntax-Fehler für den aktuell untersuchten Befehl zurückgeliefert hat, wird der betreffende Befehl der studentischen Lösung im Gesamtergebnis unmittelbar als "inkorrekt" bewertet. Liegt für das zu untersuchende SQL-Statement keine Fehlermeldung des Interpreters vor, wird dieser Teil der Einreichung mit dem korrespondierenden Befehl der Musterlösung inhaltlich verglichen und darauf überprüft, ob zum Beispiel die Ergebnis-Tabellen einer "SELECT"-Anweisung die gleichen Zeilen besitzen oder ob die Datensätze aus den "INSERT"-Anweisungen inhaltlich den Anforderungen der Aufgabenstellung genügen. Hierfür ist jedoch, für ein positives Ergebnis, von Nöten, dass die Studierenden sich hinsichtlich der Anordnung von Attributen und Reihenfolge von "INSERT"-Statements exakt an die Spezifikationen der Aufgabenstellung halten.

(7) Wurde ein Fehler ermittelt, erzeugt das Skript eine Fehlermeldung, die einen Hinweis darauf enthält, in welchem SQL-Statement ein Syntaxfehler bzw. ein Fehler inhaltlicher Natur aufgetreten ist. Um Ketten von Fehlermeldungen zu vermeiden, deren Ursprung in einem Fehler zu Beginn der SQL-Statements liegt, wird momentan überprüft ob der aktuell gefundene Fehler der erste Fehler in der Einreichung ist. Nur für diesen Fall wird die entsprechende Meldung erzeugt. Der erzeugte Hinweis beinhaltet Anmerkungen zu häufig gemachten Fehlern in dem betreffenden Statement bzw. dem betroffenen Befehlstyp. Entspricht ein Befehl auch inhaltlich dem betreffenden Befehl der Musterlösung, so wird dieser als "korrekt" bewertet.

(8) Das Ergebnis des Befehlsvergleichs wird zunächst skriptseitig zur späteren Rückmeldung an LON-CAPA verwaltet.

Wurden alle Befehle aus Musterlösung und studentischer Einreichung verglichen, wird vom Skript ein Feedback mit den Bewertungen der einzelnen Befehle erzeugt, die vom Praktomat an LON-CAPA zurückgegeben werden, so dass sie dort dem Studenten zur Verfügung gestellt werden können. Das Feedback sorgt genau dann für eine positive Bewertung der Aufgabe in LON-CAPA, wenn für keinen Befehl ein Fehler ermittelt wurde. Wie bei allen Aufgaben in LON-CAPA wird das Ergebnis, das durch diese Rückmeldung erzeugt wird ebenso in die Aufgabenstatistik des Kurses bzw. der spezifischen Aufgabe übernommen und eingepflegt, um auch dem Lehrenden die Möglichkeit des Einblicks zu geben.

Aufgrund der Tatsache, dass die Funktionsweise des hier vorgestellten Skript-Checkers unabhängig von der Konkretisierung der gestellten Datenbankaufgaben ist, lässt sich hierdurch somit eine automatische Bewertung beliebiger SQL-Aufgaben erreichen.

4 Ausblick

Das System ist seit zwei Semestern erfolgreich im Kurs Datenbanken im Einsatz gewesen und wurde frequentiert von den Studierenden eingesetzt. Als Herausforderung kann die sorgsame Spezifikation von Online-Aufgaben genannt werden um Mehrdeutigkeiten bei der Interpretation der Aufgabenstellung auszuschließen oder diese zumindest in der Musterlösung zu berücksichtigen. Durch die verbesserte Wiederverwendung wird sich der dadurch erhöhte Aufwand langfristig auszahlen.

Die Autoren danken dem Zentrum für erfolgreiches Lehren und Lernen www.ostfalia.de/zell. Dieses Vorhaben wird aus Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01PL11066H gefördert. Die

Verantwortung für den Inhalt dieser Veröffentlichung liegt beim Autor.

Literatur

[1] *J. Krinke, M. Störzer, A. Zeller*: Web-basierte Programmierpraktika mit Praktomat. In: *Softwaretechnik-Trends* 22, 3 10/2002. <http://www.st.cs.uni-sb.de/publications/details/krinke-gitrends-2002/>.

[2] *A. Stöcker, T. Chukhlova, S. Tjettmers, S. Becker, O. Bott*: E-Prüfungen mit dem LMS Moodle: Ergebnisse einer Pilotstudie. In: *Goltz et al, Informatik 2012*, 42. Jahrestagung der Gesellschaft für Informatik, GI-Edition, *Lecture Notes in Informatics*, P-208, 2012, 1808-1821.