

# **Intelligent Tutoring System (ITS) zur emotionalen Lernunterstützung**

Forschungsprojekt B

an der

Hochschule für Technik und Wirtschaft Berlin  
Fachbereich Wirtschaftswissenschaften II  
Studiengang Angewandte Informatik

Prüfer

Prof. Dr. Albrecht Fortenbacher

***Eingereicht von Max Oehme***

18. April 2021

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Emotionale Lernunterstützung in Betty's Brain</b>	<b>2</b>
2.1	Erkennung emotionaler Lernzustände . . . . .	2
2.1.1	Emotional Questionnaire . . . . .	2
2.1.2	Gesichtserkennung . . . . .	3
2.1.3	Gazedetection . . . . .	4
2.1.4	Physiologische Sensordaten . . . . .	4
2.2	Emotionales Feedback . . . . .	5
2.2.1	Textuelle Dialog . . . . .	5
2.2.2	Avatare . . . . .	6
2.2.3	Rolle von Betty . . . . .	6
2.2.4	Rolle von Mr.Davis . . . . .	7
<b>3</b>	<b>Analyse von Realisierungsmöglichkeiten</b>	<b>8</b>
3.1	Methoden zur Erkennung des emotionalen Lernzustands . . . . .	8
3.1.1	Gesichtserkennung . . . . .	8
3.1.2	Gazedetection . . . . .	9
3.1.3	Physiologische Sensordaten . . . . .	9
3.2	Möglichkeiten von emotionalem Feedback in Betty's Brain . . . . .	10
3.2.1	Textuelle Dialog . . . . .	11
3.2.2	Empathischer Avatare . . . . .	12
3.2.3	Rolle von Mr.Davis . . . . .	13
3.2.4	Rolle von Betty . . . . .	13
<b>4</b>	<b>Umsetzung</b>	<b>14</b>
4.1	Diskussion von Umsetzungsmöglichkeiten . . . . .	14

4.2 Softwareentwurf . . . . .	15
4.3 Implementierung . . . . .	17
<b>5 Diskussion</b>	<b>19</b>
<b>Literaturverzeichnis</b>	<b>20</b>

# 1 Einleitung

Die Schule ist in der aktuellen Gesellschaft der wichtigste Ort an dem sich sowohl Kinder als auch Erwachsene Wissen aneignen. Mehrere Schüler sitzen dafür klassischerweise zusammen in einem Klassenraum und werden von einem Lehrer unterrichtet. Diese Methode, die bereits seit hunderten Jahren Anwendung findet, stößt seit der COVID-19 Krise jedoch an ihre Grenzen. Schulen werden geschlossen oder operieren nur eingeschränkt mit strengen Hygiene Auflagen. Distanzlernen, also das Lernen von zu Hause über den Computer mit Webcam, ist zum Alltag für viele Schüler und Studenten weltweit geworden. Durch das eigenständige Lernen fehlt jedoch eine wichtige Komponente, der Lehrer. Ein Lehrer vermittelt nicht nur das zu Lernende, sondern unterstützt den Lernenden auch. Ohne einen Lehrer oder Tutor ist eine höhere Selbstdisziplin von den Schülern gefordert und um Fragen zu beantworten ist meist niemand qualifiziertes anwesend.

Ein Intelligent Tutoring System soll Studenten beim eigenständigen Lernen helfen. Diese Systeme nutzen simulieren einen Tutor/Lehrer der einem das Wissen vermittelt und Hilfestellung leistet. Die Emotionen des Lernenden werden vom Tutor häufig aber nicht betrachtet. Dieser Aspekt wurde im vorangehenden Forschungsprojekt A untersucht. Es wurden dazu verschiedene ITS untersucht und welche Maßnahmen für emotionelle Lernunterstützung diese nutzen. Motivierendes Feedback wird häufig umgesetzt. Es wurden die Einflüsse von Techniken des Affective Computing in Lernumgebungen betrachtet und wie der Lernerfolg gesteigert werden kann. In Bezug auf das ITS Betty's Brain konnten Vorschläge gemacht werden, welche Methoden für emotionales Lernsupport eingesetzt werden könnten.

In dieser Arbeit soll untersucht werden welche Methoden für sowohl Emotionserkennung als auch emotionales Feedback gibt es und wie diese in Betty's Brain umgesetzt werden können.

## 2 Emotionale Lernunterstützung in Betty's Brain

### 2.1 Erkennung emotionaler Lernzustände

Für gezielte emotionale Lernunterstützung ist es notwendig den Lernzustand des Lernenden zu kennen. Hat ein Tutor diese Daten kann er dem Lernenden passende Hilfestellung leisten, z.B. mit motivierendem Feedback. Solche Hilfestellungen würden nur angeboten werden wenn der passende Lernzustand erkannt wurde. Dafür muss es Methoden geben den emotionales Lernzustand zu erkennen. Ein solcher Lernzustand<sup>1</sup> sind spezielle Emotionen basierend auf Ekmans 6 Basisemotionen. Es sind Emotionen die speziell in Lernsituationen auftreten, z.B. Unterricht in der Schule, Lernen zu Hause oder Schulaufgaben bearbeiten. [1] Solche Emotionen sind u.a. Frustration, Langeweile, Neugierde, Stress und Aufmerksamkeit. Im folgenden wurden Methoden zur Emotionerkennung untersucht die in Studien mit Betty's Brain eingesetzt wurden.

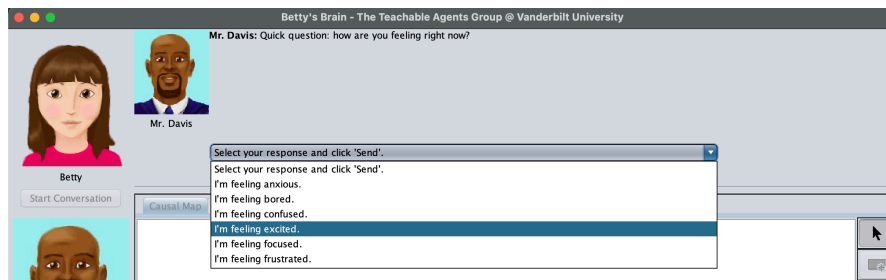
#### 2.1.1 Emotional Questionnaire

Die Erkennung von Emotionen lässt sich bereits ohne die Nutzung von Equipment wie Sensoren oder Kameras bewerkstelligen. Betty's Brain hat an sich nicht die Möglichkeit Emotionen bzw. Lernzustände automatisiert zu erkennen. Über einen Emotional Questionnaire kann man aber ein Self-Assessment zum aktuellen Lernzustand des Lernenden erhalten. Mit Hilfe der Dialogbox wird dem Nutzer alle zehn Minuten von Mr. Davis eine Frage gestellt: "How are you feeling?". Der Lernende kann darauf

---

<sup>1</sup>engl. Academic Emotion

Abbildung 2.1: Emotional Questionnaire Dialog mit Mr. Davis.



mit einer von sechs Auswahlmöglichkeiten antworten (Abbildung 2.1), die je einen Lernzustand widerspiegeln: anxious, bored, confused, excited, focused, frustrated.

### 2.1.2 Gesichtserkennung

In der Affective Computing und Bildungs Forschung wird bis heute häufig auf Facial Recognition gesetzt, die auch in aktuellen Untersuchungen zu Betty's Brain angewandt wird [2, 3, 4]. Studien benötigen dazu kein spezielles Equipment und können mit einfachen Laptop Webcams durchgeführt werden. In der Studie von [2] erhielten die Partizipanten einfache Laptops auf denen Betty's Brain installiert war. Es wurde untersucht inwiefern Basis-Emotionen mit die komplexeren akademischen Emotionen, den Lernzuständen, verknüpft sind. In weitere Studien [3] wurde die Verbindung von Gesichts Features zu Gaze Features untersucht. Es wurden mehrere Studien analysiert, bei denen während der Nutzung von Betty's Brain neben der Aufzeichnung des Gesichts über die Webcam auch Gazetracking über einen Eye-Tracker statt fand. Die Studien zeigten eine hohe Übereinstimmung der Ergebnisse aus beiden Datensätzen, von Gaze und Gesicht Features. Auch außerhalb von von Betty'S Brain zeigt der Einsatz von Facial Recognition zu Bestimmung der Emotionen gute Ergebnisse [4]. Hervorzuheben ist der Mobiliäts Aspekt bei Gesichtserkennung, da lediglich ein Laptop mit einer eingebauten Webcam benötigt wird und keine externe Hardware.

### 2.1.3 Gazedetection

Darüber hinaus gibt es Untersuchungen mittels Gesichtserkennung die Emotionen eines Lernenden zu bestimmen. Im Zusammenhang mit Betty's Brain wurde von [5] die Verwendung von Eye-Tracking untersucht. Bei diesem Verfahren wird erkannt auf welchen Bereich vom Bildschirm der Blick des Lernenden fällt [6, 5]. Es lässt sich dadurch bestimmen auf welchen Teil der Anwendung der Nutzer sich fokussiert und wie das Leseverhalten des Lernenden ist. In verschiedenen Studien wurde der Einsatz dieser Sensoren im Kontext der ITS untersucht. GazeTutor [6] sollte mit Hilfe der Daten vom Eye-Tracker die Aufmerksamkeit des Lernenden verbessern. Das Wegschauen vom Bildschirm für eine längere Zeit wurde dabei als Langeweile und mangelnde Aufmerksamkeit angenommen. Um die Aufmerksamkeit dann wieder auf die Aufgaben zu lenken, wurde der Lernende mit Dialog Nachrichten darauf hingewiesen sich auch den Tutor zu fokussieren. Die Studie ergab, dass mehr als zwei drittel der Probanden weniger als vier Interventionen erhielten. Außerdem dass besonders bei komplexen Aufgaben die das Lernverständnis fordern ein Gaze-Reactive Tutor die Lernergebnisse verbesserte. Beim MetaTutor hatten [7] Machine Learning Algorithmen eingesetzt um Langeweile und Neugierde zu erkennen und dabei Genauigkeiten von 69% und 73% erreicht.

### 2.1.4 Physiologische Sensordaten

Sich verändernde Gehirnaktivität, Herzrate und Schweiß sind Auswirkungen durch verschiedenen Emotionen die sich mit physiologischen Sensoren messen lassen. Im Gegensatz zu Gesicht-Features oder Eye-Gaze werden diese Eigenschaften vom Autonomen Nervensystem (ANS) gesteuert, das großteils unwillkürlich aktiv ist und nur schwer zu kontrollieren [8, 9]. Bei den physiologischen Signalen gibt es viele unterschiedliche Arten, u.a. Elektroenzephalograf (EEG), Elektro-Kardiogramm (EKG), Skin Temperature (SKT), Galvanic Skin Response (GSR) und Elektromyogram (EMG). In Intelligent Tutoring Systems wie Betty's Brain wird der Einsatz von physiologischen Sensoren zur Emotionserkennung häufig nicht betrachtet [10]. In einer kürzlich Veröffentlichten Studie von Alqahtani et al. [10] sollten 27 Personen einen Online Englisch Sprachtest durchführen. Während des Tests trugen die Partizipanten Sensoren für EEG, EKG, und EMG. Die Ergebnisse zeigten das der Einsatz dieser

Sensoren in einem ITS das fehlen eines menschlichen Tutors ausgleicht, da durch die Sensordaten es möglich wird den Affective State des Lernenden zu verstehen und sich ihm anzupassen. Auch Studien mit anderen ITS, wie Autoturo [11] und MetaTutor [12], haben den Einsatz von physiologischen Sensordaten untersucht. Diese zeigten das mit diesen Sensoren hohe Genauigkeiten bei der Erkennung erzielt werden können, jedoch sollten Sensoren eher in Kombination mit anderen Datenquellen verwendet werden.

## **2.2 Emotionales Feedback**

Bis auf beim Emotional Questionnaire spielen Emotionen für das Feedback bisher noch keine Rolle. Das Feedback lässt sich aber um eine emotionale Komponente erweitern. Motivierende Nachrichten die bisher noch nicht den Lernzustand des Nutzers in Betracht ziehen, können dadurch emotional zielgerichteter motivieren. Zum Beispiel, die nach einem abgeschlossenen Quiz angezeigten Nachrichten von Mr.Davis und Betty werden bisher nur vom Ergebnis des Quizes beeinflusst. Hier könnten zusätzlich Informationen über den Lernzustand des Nutzers einbezogen werden.

### **2.2.1 Textuelle Dialog**

Feedback wird in Betty's Brain nur in Text über die Dialoge mit den Avataren wiedergegeben. An verschiedenen Stellen wird dem Nutzer Feedback gezeigt, z.B. nach der Auswertung des Quizes, wenn man Betty's Wissensstand abfragt oder auch wenn man Mr.Davis um Hilfe bittet. Es werden unterschiedliche Nachrichten angezeigt, die den Nutzer motivieren sollen. Festgestellte Emotionen, z.B. durch den Emotional Questionnaire, haben hier jedoch noch keinen Einfluss auf den Inhalt der Nachricht. Im Gegensatz dazu werden beim Emotional Questionnaire die Emotionen für das Feedback mit einbezogen. Je nachdem mit welchem Lernzustand, bzw. Emotion, der Lernende antwortet, werden ihm andere Nachrichten von Mr.Davis in der Dialogbox angezeigt.



### 2.2.2 Avatare

Im Forschungsprojekt A wurde bereits die Möglichkeit untersucht, die beiden Avatare für Feedback zu verwenden. In Betty's Brain sind die Avatare Mr. Davis und Betty statische Bilder, die im Fall von Betty auch durch Anpassungen personalisiert werden können. Eine Nutzung zur Wiedergabe von Feedback ist nicht implementiert. Die Anwendung kann aber abgeändert werden, sodass jeder der Avatare eine Auswahl an verschiedenen Bildern anzeigen kann. Diese Bilder zeigen unterschiedliche Gesichtsausdrücke die für eine Darstellung von simulierten Emotionen des Avatars dienen sollen. Die Anzeige der Ausdrücke wird dann an das Text-Feedback in den Dialogen gekoppelt. Entsprechend welcher Art von Feedback im Dialog vorkommt werden andere Gesichter gezeigt. Kommt von Mr. Davis eine motivierende Nachricht, so wird der Avatar von ihm mit einem positiven freundlichen Gesichtsausdruck ausgestattet. Mr. Davis und Betty können dabei ein unterschiedliches Verhalten haben, entsprechend ihrer ebenso unterschiedlichen Rollen in der Anwendung. Mr. Davis als Supervisor gibt Feedback und Motivation zu den Aufgaben die man bearbeitet. Betty auf der anderen Seite ist eine Reflexion des Lernenden. Dem entsprechend kann dann der Avatar den eigenen Lernzustand widerspiegeln.

### 2.2.3 Rolle von Betty

Betty ist die Hauptfigur in der Anwendung und den Schüler den der Nutzer zu unterrichten hat. Als Nutzer muss man Betty den Inhalt der Wissensbasis lehren indem man die verschiedenen Konzepte in einer *concept map* miteinander verknüpft. Während des Lernprozesses kann man Betty nach einzelnen Konzepten fragen, um die Korrektheit zu prüfen, oder man lässt sie ein Quiz durchführen. Betty nutzt vereinfachtes qualitatives Reasoning [13] um die einzelnen Konzepte zu verknüpfen und Antworten zu generieren. Anhand von Betty's Leistung in einem Quiz kann der Nutzer erkennen welche Teile der *concept map* überarbeitet werden müssen. Dazu kann Betty auch auf den Antworten basierende Erläuterungen bieten, die dem Nutzer helfen sollen die Map zu analysieren und Fehler zu beseitigen. Zusätzlich bietet Betty nach einem Quiz motivierendes Feedback durch den Ausdruck von Freude wenn ihre Ergebnisse sich verbesserten oder Enttäuschung wenn die Ergebnisse schlechter wurden. [13]

#### **2.2.4 Rolle von Mr.Davis**

Der Agent Mr. Davis ist für Betty ein Mentor und für den Nutzer ein Superwiser. Bei einem Quiz stellt er die Aufgaben und nach Abschluss bietet Feedback, indem er Strategien fürs Lernen vorschlägt, wenn die Ergebnisse nicht so gut ausfielen. Auch während des Lernens kann man mit Mr. Davis interagieren und ihn Fragen rund um die Anwendung stellen, z.B. wie man eine *concept map* erstellt oder die Quizes funktionieren. Abgesehen davon stellt Mr. Davis in einem zehn Minuten Abstand regelmäßig den Emotional Questionnaire. [13]

## 3 Analyse von Realisierungsmöglichkeiten

### 3.1 Methoden zur Erkennung des emotionalen Lernzustands

In dieser Sektion sollten Möglichkeiten analysiert werden Emotionerkennung in Betty's Brain einzubinden.

#### 3.1.1 Gesichtserkennung

Für die Erkennung von Emotionen auf Gesichtsfeatures wird keine spezielle technische Ausrüstung benötigt, eine einfache Webcam ist dafür ausreichend [14]. Für eine Klassifizierung in Lernzustände werden unterschiedliche Gesichtsfeatures untersucht, die über die Webcam aufgezeichnet wurden. Die verschiedenen Herangehensweisen unterscheiden sich vor allem dabei welche Emotionen erkannt werden sollen und welche Gesichtsfeatures dabei einbezogen werden. In [14] wurden die Abstände zwischen festgelegten Stellen im Gesicht berechnet, im Vergleich zu einem Gesicht im neutralen Zustand. Die Abstände wurden dann den sechs Basis Emotionen von Ekman zugeordnet. Bei den Arbeiten von [15] wurden für die Pattern Recognition der distinkten Gesichtsfeatures Singular Value Decomposition (SVD) und zur Klassifizierung Hidden Markov Model (HMM) verwendet. In diesem Fall wurden jedoch direkt sechs Lernzustände und nicht die Basis Emotionen bestimmt. Lernzustände bzw. "Academic Emotions" [1] sind ein besserer Indikator des Lernerfolgs als die Basis Emotionen die in [14] und [16] angewandt wurden [1]. Bei den Arbeiten von [15], [17] und [18], die eine Klassifizierung in Lernzustände vornahmen, konnte vor allem Engagement mithilfe von Gesichtserkennung gute Klassifizierungsergebnisse erzielen während Frustration schwieriger zu erkennen war.

Wie bereits in vorherigen Studien zu Betty's Brain kann Gesichtserkennung wieder mit dem Einsatz der eingebauten Webcam, des Laptops auf dem die Anwendung

läuft, eingesetzt werden [2, 3, 4].

### **3.1.2 Gazedetection**

Abschweifende Gedanken haben einen signifikant negativen Einfluss auf den Lernerfolg [19]. Mithilfe von Eye-Tracking Daten lässt sich erkennen wann der Lernende mit seinen Gedanken "woanders" ist. Dies lässt sich mit Lernzuständen wie Langeweile, Neugierde und Aufmerksamkeit in Verbindung bringen [7, 19]. Festgestellt wurden dieses Verhalten vor allem bei Aufgaben bei denen der Lernende viel eigenständig lesen muss. Zum Ende einer Seite war dies verstärkt, was auf eine Reflektieren des Nutzers zum Gelesenen zurückzuführen ist [19]. Durch das selbstständige Ausarbeiten der Lösung aus einer Wissensbasis wird man dieses Verhalten auch bei Nutzern von Betty's Brain nachweisen können. Die Wissensbasis ähnelt einem elektronische Buch, bei dem der Inhalt nur in Kapitel und nicht nach Seiten gegliedert ist. Mit einem Eye-Tracker kann man die Qualität des Lesens beurteilen, indem man feststellen kann wie die Aufmerksamkeit beim Lesen kritischer Abschnitte war. Betty's Brain kann dem Nutzer dann bitten die vernachlässigten Abschnitte erneut zu lesen oder weitere Aufgaben zum Vertiefen anbieten [19].

Durch den Eye-Tracker kann man generell ermitteln auf welche Bereiche, Schaltflächen, Textinhalte und weiteres der Nutzer blickt. Das ermöglicht es dem Nutzer visuelles Feedback nicht nur während des Lesen zu erhalten, sondern auch in anderen Situationen, z.B. bei Schwierigkeiten im Umgang mit dem Userinterface. Ebenfalls lassen sich Probleme mit der Anwendung identifizieren, zum Beispiel Aufgaben bei denen der Lernende Probleme mit der Bearbeitung hat oder die Langeweile oder Frustration verursachen [6].

### **3.1.3 Physiologische Sensordaten**

Bei dem Einsatz von physiologische Sensordaten zur Erkennung von Lernzuständen wird je nachdem welche Art von Signaltyp benötigt wird, entsprechend anderes Equipment verwendet. Beim EEG werden Sensoren am Kopf, bei EMG verschiedene Sensoren an ausgewählten Muskeln und bei Signalen wie EDA, Herzrate oder Hauttemperatur Sensoren an einem Ort wie das Handgelenk benötigt.

Mit der Nutzung EEG lassen sich die Basis Emotionen akkurat bestimmen und die Klassifizierung von Signalen in Emotionen erreicht mit Machine Learning Algorithmen wie Support-Vector-Machine eine hohe Genauigkeit [20, 21]. Bei der Erkennung von Lernzuständen zeichnen sich jedoch schlechtere Ergebnisse bei der Klassifizierung ab [22]. Außerdem bestehen Probleme bei dem Einsatz vom EEG außerhalb von Laborbedingungen. Zum einen ist das Equipment teuer in der Anschaffung und zum anderen sind die Signale schwächer als bei anderen Methoden was sie Anfällig für Störungen machen [20].

Bei einem EMG werden Elektroden direkt auf die Haut aufgetragen oder mit einer Nadel injiziert. Für die Bestimmung von Emotionen werden hierfür meist Gesichtsmuskeln genommen, da diese emotionalen Reaktionen am stärksten abbilden [21]. Während die Genauigkeiten einer Klassifizierung mit SVM recht hoch sind, werden EMG zu Emotionsbestimmung selten eingesetzt, da sie ähnliche Probleme haben wie EEG. Die Anbringung der Elektroden beeinflusst den Komfort der Person und bringt Limitierungen mit sich. Außerdem ist das EMG nicht so zuverlässig, da es nur stärkere Emotionen erkennen kann, denn nur diese bringen auch eine Reaktion in den Gesichtsmuskeln zum Vorschein. Ähnlich wie das EEG leidet auch das EMG an Interferenzen durch Umgebungsbedingungen wie Sonnenlicht oder Elektrostatik, aber auch ungewollte Muskelbewegungen verfälschen das Ergebnis [21].

Im Gegensatz zum EEG und EMG können Sensoren zur Bestimmung von Herzrate, Hauttemperatur und Hautleitfähigkeit einfach in Form eines Armbands am Handgelenk angebracht werden, ähnlich wie es mit kommerziell erhältlichen Fitness Armbändern der Fall ist. Sensoren für diese Daten behindern den Nutzer nicht und verringern auch nicht den Komfort [9].

### **3.2 Möglichkeiten von emotionalem Feedback in Betty's Brain**

In vielen ITS wurden unterschiedliche Methoden für verschiedene Arten von Feedback untersucht und umgesetzt, jedoch sind nicht alle dieser Methoden auch in Betty's Brain anwendbar. In der folgenden Sektion wurden Methoden für Feedback untersucht die in Betty's Brain eingesetzt werden können.

### 3.2.1 Textuelle Dialog

Über die Dialogbox in Betty's Brain werden bereits motivierende Nachrichten an den Lernenden gesendet. Jedoch ist die Interaktion mit Betty und Mr.Davis limitiert auf Konversationen die vom Nutzer aus gestartet werden oder Feedback das nach Aktionen wie dem Quiz gegeben wird. In einer Studie von [23] wurde in dem ITS MetaTutor der Einfluss von Prompts während dem Lesen untersucht. MetaTutor bietet wie Betty's Brain einen Pädagogischen Agent mit einem Avatar, eine Wissensbasis und die Möglichkeit mit dem PA zu interagieren. Während der Lernende in der Wissensbasis gelesen hat, wurden Prompts mit Feedback angezeigt. Wann ein Prompt angezeigt hing dabei von Faktoren wie Interaktion, Zeit die man bereits liest, Anzahl der besuchten Seiten usw. ab. Die Studie, die mit 69 Bachelor Studenten durchgeführt wurde, ergab das sich die Leistung der Lernenden verbessert hat. Die Studenten brauchten außerdem weniger Zeit beim Lernen und haben weniger Seiten besucht. Material wie Diagramme wurden genauer untersucht und mit einzelnen Aufgaben wurde sich intensiver beschäftigt. Durch die Ähnlichkeit in diesen Bereichen kann diese Methode auch in Betty's Brain vergleichbar umgesetzt werden.

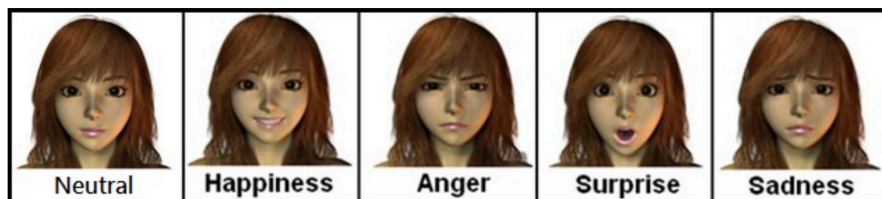
Einige ITS setzen bei Interaktion und Feedback auf Dialoge in natürlicher Sprache [24], wie AutoTutor [25] und GazeTutor [6]. In einem Review von Palafines et al. [24] wurden 33 ITS die natürliche Sprache verwenden untersucht. Die Dialoge wurden in den ITS unterschiedlich umgesetzt, jedoch bilden sich Gemeinsamkeiten ab. Generell dient der Dialog als Methode um vom Tutor Instruktionen zu erhalten und Lösungen für die Aufgaben vorzuschlagen. Der Lernende kann zum Beispiel Rückfragen zu einer vorher gestellten Frage stellen und um weitere Erläuterungen bitten, der Tutor liefert dem Lernenden daraufhin zusätzliche Hinweise zur Bewältigung der Aufgabe. Macht der Lernende Fehler oder erkennt der Tutor das der Lernende einen falschen Ansatz verfolgt, kann der Tutor auch hier Hinweise geben. Im Umgang mit falschen Antworten geben die meisten Tutoren dem Nutzer entweder die Lösung oder animieren den Lernenden die Aufgabe erneut zu bearbeiten, mit einigen Hinweisen vom Tutor. Je nach Leistung des Lernenden können nicht nur Hinweise zur Lösungsfindung gegeben werden, sondern auch Aufforderungen bestimmte Themen erneut nachzuschlagen und Hervorheben das Teile der Antwort inkorrekt sind oder fehlen. Abgesehen davon erhalten Nutzer in den untersuchten ITS auch Feedback während

des Dialogs. AutoTutor bietet dabei die Möglichkeit anhand des Lernzustands dem Nutzer Feedback zu geben, das den Lernenden zum Beispiel motivieren soll [25]. Die Einbindung von Feedback und Hilfen bei den Aufgaben als Dialog in natürlicher Sprachen lässt den Tutor menschlicher wirken und kann den Lernerfolg verbessern.

### 3.2.2 Empathischer Avatare

Durch die Darstellung des pädagogischen Agent als anthropomorphen Avatar wie mit Betty und Mr. Davis soll die PA menschlicher wirken lassen und Lernende sollen eine empathische Verbindung mit ihnen knüpfen. Avatare mit menschlichem Aussehen und Verhalten haben einen positiven Einfluss auf die Ergebnisse in Lernsituationen [26, 27]. Mit einem empathischen Avatar ist der PA erfolgreicher damit den Lernenden zu überzeugen mehr Aufgaben zu bearbeiten. Lernenden die Aufgaben von einem empathischen Avatar erhalten, wollen diesen nicht enttäuschen und versuchen eher dessen Erwartungen zu erfüllen [26]. Das Design eines empathischen Avatars

Abbildung 3.1: Darstellung von emotionalen Zuständen in 3D Avatar. [26]



beinhaltet die Darstellung von Gesichtsausdrücken, Handgesten, Körperbewegungen und eine Stimme [26]. Empathische Reaktionen basieren auf Aktionen in der Anwendung (z.B. Starten der Anwendung) und auf den emotionalen Lernzustand des Lernenden. Das Gesicht des Avatars zeigt dem entsprechend unterschiedliche Ausdrücke, die die sechs Basis Emotionen von Ekman darstellen (Abbildung 3.1), sowie weitere Emotionen die in einer Lernanwendung nützlich sind (z.B. Beunruhigung und Neutral). Solche Reaktionen sind ein beunruhigtes Gesicht bei Langeweile oder "Happiness" wenn "Confidence" erkannt wurde. [26]

### **3.2.3 Rolle von Mr.Davis**

In der Rolle eines Mentors ist es die Aufgabe des pädagogischen Agent "Mr. Davis" dem Lernenden zu helfen [13]. Mr. Davis bietet darüber hinaus auch Lösungsstrategien für Probleme die der Lernende nicht lösen konnte. Mit der Möglichkeit Emotionen zu erkennen kann Mr. Davis bereits früher intervenieren und während der Bearbeitung der Aufgaben Unterstützung anbieten. Erkennt er das der Lernende bei der Bearbeitung einer Aufgabe Frustriert ist, so kann dieser Hinweise geben die dem Nutzer helfen. Mr. Davis als Mentor soll den Lernenden gegebenenfalls auch dazu bringen mehr Aufgaben zu bearbeiten. Besonders wenn eine niedrige Motivation oder Langeweile erkannt wird, soll Mr. Davis den Lernenden davon überzeugen härter zu arbeiten [26].

### **3.2.4 Rolle von Betty**

In Betty's Brain ist man der Lehrer der Betty unterrichtet, demnach ist Betty eine Repräsentation des Lernenden in der Anwendung. Der Avatar lässt sich darüber hinaus vollständig anpassen was es dem Nutzer ermöglicht sich selbst so gut es geht darzustellen. Studien haben ergeben das die Selbst-Repräsentation in einer virtuellen Umgebung einen starken Einfluss auf das Verhalten [27]. Verhalten das von dem virtuellen Abbild präsentiert wird, beeinflusst das eigene Verhalten in gleicher Weise. So wird jemand der zum Beispiel sich sportlich nicht betätigt und dann das virtuelle Selbst sieht wie es auf einem Laufband läuft, dazu angetrieben mehr sportlichen Aktivitäten nachzugehen.

Der pädagogische Agent Betty und ihr Avatar können zur Verbesserung des Lernverhaltens zum einen den eigenen Lernzustand anzeigen und gegebenenfalls den wünschenswerten Lernzustand darstellen. Ist der Lernende mit der Aufgabe überfordert und frustriert kann Betty dies zeigen, aber daraufhin dem Nutzer einen fröhlicheren Zustand darstellen. Dazu sollte Betty entsprechen motivierende Nachrichten ausgeben können. Dies soll den Lernenden von denen Emotionen abbringen, die sich negativ auf den Lernerfolg auswirken (z.B. Frustration, Langeweile etc.), und den Nutzer motivieren.



## 4 Umsetzung

### 4.1 Diskussion von Umsetzungsmöglichkeiten

Für eine Umsetzung in der Anwendung Betty's Brain muss entschieden werden welche der Kapitel 3 untersuchten Methoden eingesetzt werden sollen. Bei den analysierten Feedback-Möglichkeiten werden die empathischen Avatare in Verbindung mit Text-Feedback umgesetzt. Die Avatare sind eine Funktionalität die Betty's Brain bereits besitzt und die mit dynamischen sich bewegenden Avataren erweitert werden kann. Studien zeigten das der Einsatz solcher Avatare in Lernumgebungen von Vorteil ist [26]. Auf die Nutzung von 3D Avataren wird aber vorerst verzichtet, da Betty's Brain nicht die Möglichkeit besitzt 3D Grafiken oder Inhalte anzuzeigen. Parallel zu den Avataren kann Text Feedback in der Dialogbox angezeigt werden, wodurch bestehende Funktionalität erweitert wird. Dialoge in Natürlicher Sprache kommen jedoch nicht zum Einsatz. Solche Dialoge erwarten ein System das Natürliche Sprache generieren kann, welches Betty's Brain nicht ist. Mit diesen Möglichkeiten der Anzeige wird den pädagogischen Agents die Funktionalität erweitert. Betty's wird über den Avatar periodisch den emotionalen Lernzustand des Nutzers darstellen [27]. Beim überschreiten von Grenzwerten erhält der Nutzer ein Feedback von Mr. Davis das motivieren soll [26].

Damit das Feedback an den Avataren und als Text angezeigt werden soll, benötigt die Anwendung Daten zu den emotionalen Lernzuständen. Methoden wie Gesichtserkennung und Gazedetection sind geeignet, da es viel Forschung dazu gibt und auch das Equipment günstig und im Handel erhältlich ist [14]. Besonders Gazedetection hat weitere Möglichkeiten im Lernkontext, da man nicht nur den emotionellen Lernzustand erkennt, sondern auch das Leseverhalten analysieren kann [19]. Für die Nutzung von physiologischen Sensordaten benötigt man spezielles Equipment wie ein EEG oder EMG. Dieses Equipment ist teuer und sehr empfindlich für Störungen,

weswegen diese unter Laborbedingungen eingesetzt werden [20, 21]. Sensorarmbänder sind hier geeigneter, da sie komfortabler für den Nutzer sind und robuster gegen Störungen [9].

Für die Implementierung werden jedoch nur die genannten Feedback Methoden umgesetzt, da für die Emotionserkennung zum einen das nötige Equipment nicht vorhanden ist und die Umsetzung sehr aufwendig ist. Die Daten sollen daher nur simuliert werden.

## 4.2 Softwareentwurf

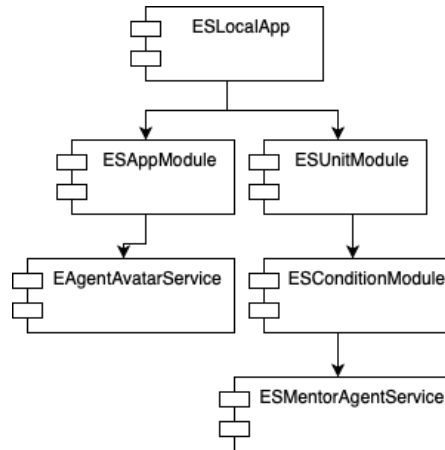
Die Anwendung Betty's Brain ist modular aufgebaut, wodurch ermöglicht wird angepasste Versionen der Anwendung bereitzustellen die neue Funktionalitäten bieten oder Bestehende zu modifizieren. Für die Entwicklung eines Moduls, das Funktionalitäten für emotionale Lernunterstützung bietet, wurde der Quellcode von Betty's Brain analysiert. Das System besteht aus 13 einzelnen Modulen die teilweise voneinander abhängig sind. Zu diesen Modulen gehören ein Kernmodul "corelib", das die Grundfunktionalitäten bereitstellt, Module für die pädagogischen Agents, die Casual Map, das Quiz und die Wissensbasis. Diese Module werden alle in dem App-Modul "Bettysbrain-app" vereint, das u.a. auch das Userinterface bereitstellt. Die vorgeschlagenen Umsetzungsmöglichkeiten betreffen hauptsächlich das Userinterface und werden daher im App-Modul Anwendung finden.

Zur Modulentwicklung in Betty's Brain werden zunächst Grundmodule benötigt:

- *launcher* der der Einstiegspunkt ist und die Basis Anwendungsfunktionen einbindet und die Lerneinheit *unit* einbindet
- *unit* zum laden der Inhalte wie der Wissensbasis
- *condition* um den Service der das Verhalten des Pädagogischen Agent bestimmt einzubinden

Für eine Version von Betty's Brain die Emotional Support umfasst müssen diese Submodule als entsprechende Emotional Support Variante implementiert werden, damit die angepasste Funktionalität des pädagogischen Agents greifen kann. Die

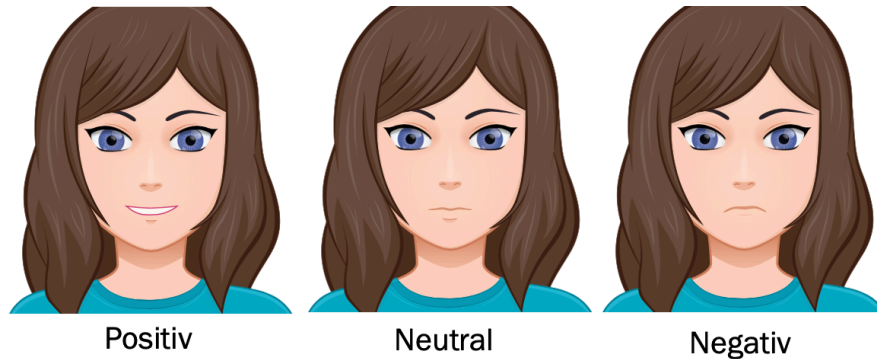
Abbildung 4.1: Benötigte Submodule für Entwurf eines Emotional Support Betty's Brain.



Funktionalitäten die der Agent bekommt werden in einem *MentorAgentService* definiert und geladen, der von der *condition* eingebunden wird. Es ergibt sich eine Modul-Hierarchie in der der *launcher* "ESLocalApp" als Einstiegspunkt zuerst geladen wird und der "ESMentorAgentService" zum Schluss (Abbildung 4.1). Ein Agent Service wird zum größten Teil von *sensors* und *behaviours* gesteuert, die entsprechend auf Events in der Anwendung (z.B. Buttons oder Zeitsteuerung) reagieren und darauf eine Antwort generieren (z.B. Dialogtext). Für Emotional Support würde hierfür ein *sensor* benötigt, der die Emotion Recognition Signale empfängt, und ein *behaviour* um bei Werten die über einem Grenzwert sind eine Feedback Nachricht im Dialogfeld zeigt.

Zum dynamischen Ändern der Avatare muss eine geänderte Version des *AgentAvatarService* eingebunden werden. Die Standard Variante des Service ist dafür verantwortlich die Avatare für Mr. Davis und Betty bereitzustellen. Außerdem wird der Avatar von Betty, je nach eingestellter Konfiguration vom Nutzer, aus den einzelnen Teilen des Gesichts zusammengestellt. Wird ein Avatar benötigt (z.B. beim Anzeigen des Dialogs), wird die entsprechende Methode aufgerufen und die Bilder geladen. Diese Funktion kann man sich bei einer Emotional Support Version zu nutze machen, indem der Service über einen *EventBus* auf Updates (neue Emotion erkannt) vom *sensor* des AgentService wartet und das anzuzeigende Bild von Betty tauscht. Je nach Emotionalem Zustand wird dann eines von drei Bildern geladen die

Abbildung 4.2: Avatar Varianten von Betty die verschiedene Gesichtsausdrücke zeigen.



die Zustände als "positiv", "neutral" und "negativ" zeigen (Abbildung 4.2).

### 4.3 Implementierung

Obwohl das Betty's Brain System modular aufgebaut ist stellte sich das dynamische Laden der Avatare als schwierig heraus. Die Anfrage nach dem Avatar Bild an den AvatarService wird nur gestellt, wenn ein Avatar zum ersten Mal angezeigt werden soll, z.B. beim Programmstart oder einblenden der Dialogbox. Ein verändern des Avatars würde erst sichtbar sein, sobald eine neues UI Element mit dem Avatar Bild angezeigt werden soll, z.B. das Starten eines Dialogs. Erhält der AvatarService ein Event zum Ändern des Avatar Bildes, muss diese daraufhin über den EventBus ein Event an den ViewController senden um ein Neuladen des UI auszulösen.

Das UserInterface von Betty's Brain wurde mit Java Swing entwickelt. Wird das Event zum Neuladen im ViewController ausgelöst, fragt der Controller erneut nach dem Avatar Bild aus dem AvatarService. Das Bild wird dann im entsprechenden *JLabel*, das die Avatare anzeigt, eingesetzt und es wird die *revalidate* Methode des *JFrame* ausgelöst. Letzteres bewirkt ein erneutes Rendern des UI mit gegebenenfalls veränderten Swing Objekten, wie dem veränderten *JLabel* der Avatare.

Damit unterschiedliche Avatare angezeigt werden können, muss es eine Möglichkeit geben Emotionen zu erkennen. Aktuell sollen diese Emotionen nur simuliert werden und es wird nicht mit echten Sensordaten gearbeitet. Dafür wird im Agent-

Service ein *sensor* eingebunden, der sowohl für echte Daten als auch Simulierte als Empfangsstelle dienen soll. Der *sensor* würde die analysierten Daten empfangen als erkannte Emotionen, die im Fall der Simulation zufällig generiert werden. In periodischen Abständen werden diese Daten über den EventBus an den AvatarService gesendet. Zusätzlich werden die Daten an ein *behaviour* weitergegeben, das beim überschreiten von Grenzwerten neben dem Gesichtsausdruck des Avatars auch eine Feedback Nachricht als Text anzeigt.

## 5 Diskussion

Mit einer Implementierung von emotionalem Support in Betty's Brain konnten die Möglichkeiten die das System bietet gezeigt werden. Durch die Modularität sind verschiedene Methoden für sowohl Emotionserkennung als auch emotionales Feedback umsetzbar. Bei der Erkennung von Emotionen lassen sich verschiedene Detektoren im gleichen System implementieren, wodurch diese sich gegeneinander evaluieren lassen. Für die Implementierung von Feedback ist man etwas eingeschränkter, da man mit den Begrenzungen des Userinterface arbeiten muss. Änderungen die die bestehende Grundstruktur des UI nur erweitern lassen sich auch einfache Weise umsetzen. Außerhalb von der Grundarchitektur sind durch das Modulkonzept aber auch Änderungen möglich, jedoch teilweise mit erheblichem Aufwand verbunden. So mussten für die Umsetzung der empathischen Avatare einige Grundfunktion geändert damit Avatare dynamisch geladen und erneuert werden können, da eine solche Funktion ursprünglich nicht vorgesehen war. Auch der vorgeschlagene Dialog in Natürlicher Sprache ist vom System limitiert. Da ein solches System Generierung von Natürlicher Sprache in Echtzeit verlangt. Konversationen mit den Agents sind in Betty's Brain jedoch fest definiert.

Während Studien die betrachteten Methoden evaluiert hatten und deren Wirkung als großteils positiv einschätzten, ist es notwendig die implementierten Funktionen mit Betty's Brain separat zu evaluieren. Dafür muss untersucht werden welche Gesichtsausdrücke die Avatare zu welchen Zeitpunkt zeigen, die genauen Inhalte des motivierenden Feedbacks und aber auch Grenzwerte ab wann Feedback angezeigt werden muss.

## Literaturverzeichnis

- [1] R. Pekrun and E. J. Stephens, "Academic emotions." in *APA educational psychology handbook, Vol 2: Individual differences and cultural and contextual factors*. American Psychological Association, 2012, pp. 3–31.
- [2] A. Munshi, S. Mishra, N. Zhang, L. Paquette, J. Ocumpaugh, R. Baker, and G. Biswas, "Modeling the relationships between basic and achievement emotions in computer-based learning environments," in *International Conference on Artificial Intelligence in Education*. Springer, 2020, pp. 411–422.
- [3] K. Sharma, S. Mishra, Z. Papamitsiou, A. Munshi, B. K. De, G. Biswas, and M. Giannakos, "Towards obtaining facial proxies for gaze behaviour in tel," 2020.
- [4] K. Sharma, E. Niforatos, M. Giannakos, and V. Kostakos, "Assessing cognitive performance using physiological and facial features," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 4, no. 3, pp. 1–41, sep 2020.
- [5] R. Rajendran, A. Kumar, K. E. Carter, D. T. Levin, and G. Biswas, "Predicting learning by analyzing eye-gaze data of reading behavior." *International Educational Data Mining Society*, 2018.
- [6] S. D’Mello, A. Olney, C. Williams, and P. Hays, "Gaze tutor: A gaze-reactive intelligent tutoring system," *International Journal of human-computer studies*, vol. 70, no. 5, pp. 377–398, 2012.
- [7] N. Jaques, C. Conati, J. M. Harley, and R. Azevedo, "Predicting affect from gaze data during interaction with an intelligent tutoring system," in *International Conference on Intelligent Tutoring Systems*. Springer, 2014, pp. 29–38.

- [8] L. Shu, J. Xie, M. Yang, Z. Li, Z. Li, D. Liao, X. Xu, and X. Yang, "A review of emotion recognition using physiological signals," *Sensors*, vol. 18, no. 7, p. 2074, jun 2018.
- [9] S. Wioleta, "Using physiological signals for emotion recognition," in *2013 6th International Conference on Human System Interactions (HSI)*. IEEE, jun 2013.
- [10] F. Alqahtani, S. Katsigiannis, and N. Ramzan, "Using wearable physiological sensors for affect-aware intelligent tutoring systems," *IEEE Sensors Journal*, vol. 21, no. 3, pp. 3366–3378, feb 2021.
- [11] H. C. Lane and S. K. D’Mello, "Uses of physiological monitoring in intelligent learning environments: A review of research, evidence, and technologies," in *Mind, Brain and Technology*. Springer International Publishing, dec 2018, pp. 67–86.
- [12] M. S. Hussain, O. AlZoubi, R. A. Calvo, and S. K. D’Mello, "Affect detection from multichannel physiology during learning sessions with AutoTutor," in *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2011, pp. 131–138.
- [13] G. Biswas, J. R. Segedy, and K. Bunchongchit, "From design to implementation to practice a learning by teaching system: Betty’s brain," *International Journal of Artificial Intelligence in Education*, vol. 26, no. 1, pp. 350–364, 2016.
- [14] M. B. Ammar, M. Neji, A. M. Alimi, and G. Gouardères, "The affective tutoring system," *Expert Systems with Applications*, vol. 37, no. 4, pp. 3013–3023, apr 2010.
- [15] K. O. Akputu, K. P. Seng, and Y. L. Lee, "Facial emotion recognition for intelligent tutoring environment," in *2nd International Conference on Machine Learning and Computer Science (IMLCS’2013)*, 2013, pp. 9–13.
- [16] J. Khalfallah and J. B. H. Slama, "Relevant metrics for facial expression recognition in intelligent tutoring system," in *Innovations in Smart Learning*. Springer Singapore, sep 2016, pp. 119–122.



- [17] R. Zatarain-Cabada, M. L. Barron-Estrada, F. Gonzalez-Hernandez, and H. Rodriguez-Rangel, "Building a face expression recognizer and a face expression database for an intelligent tutoring system," in *2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT)*. IEEE, jul 2017.
- [18] K. D. Sidney, S. D. Craig, B. Gholson, S. Franklin, R. Picard, and A. C. Graesser, "Integrating affect sensors in an intelligent tutoring system," in *Affective Interactions: The Computer in the Affective Loop Workshop at*, 2005, pp. 7–13.
- [19] R. Bixler and S. D'Mello, "Automatic gaze-based user-independent detection of mind wandering during computerized reading," *User Modeling and User-Adapted Interaction*, vol. 26, no. 1, pp. 33–68, 2016.
- [20] T. Xu, Y. Zhou, Z. Wang, and Y. Peng, "Learning emotions EEG-based recognition and brain activity: A survey study on BCI for intelligent tutoring system," *Procedia Computer Science*, vol. 130, pp. 376–382, 2018.
- [21] A. Dzedzickis, A. Kaklauskas, and V. Bucinskas, "Human emotion recognition: Review of sensors and methods," *Sensors*, vol. 20, no. 3, p. 592, jan 2020.
- [22] J. Azcarraga, N. Marcos, and M. T. Suarez, "Modelling eeg signals for the prediction of academic emotions," in *Workshop on Utilizing EEG Input in Intelligent Tutoring Systems (ITS2014 WSEEG)*, 2014, p. 1.
- [23] R. Azevedo, R. S. Landis, R. Feyzi-Behnagh, M. Duffy, G. Trevors, J. M. Harley, F. Bouchet, J. Burlison, M. Taub, N. Pacampara, M. Yeasin, A. K. M. M. Rahman, M. I. Tanveer, and G. Hossain, "The effectiveness of pedagogical agents' prompting and feedback in facilitating co-adapted learning with MetaTutor," in *Intelligent Tutoring Systems*. Springer Berlin Heidelberg, 2012, pp. 212–221.
- [24] J. Paladines and J. Ramirez, "A systematic literature review of intelligent tutoring systems with dialogue in natural language," *IEEE Access*, vol. 8, pp. 164 246–164 267, 2020.
- [25] A. C. Graesser, P. Chipman, B. C. Haynes, and A. Olney, "Autotutor: An intelligent tutoring system with mixed-initiative dialogue," *IEEE Transactions on Education*, vol. 48, no. 4, pp. 612–618, 2005.

- [26] G.-D. Chen, J.-H. Lee, C.-Y. Wang, P.-Y. Chao, L.-Y. Li, and T.-Y. Lee, "An empathic avatar in a computer-aided learning program to encourage and persuade learners," *Journal of Educational Technology & Society*, vol. 15, no. 2, pp. 62–72, 2012.
- [27] A. L. Baylor, "Promoting motivation with virtual agents and avatars: role of visual presence and appearance," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 364, no. 1535, pp. 3559–3565, dec 2009.