

Untersuchung Entscheidungskriterien bei den Menschen

Valéry Toda
Neuronale Informationsverarbeitung
Technische Universität Berlin
todaval@gmail.com

Betreuer: Robert Martin, PhD
Bernstein Center for Computational Neuroscience
Humboldt Universität Berlin
robert.martin@bccn-berlin.de

Kurzfassung

Im Rahmen dieses Projekts werden Entscheidungsstrategien bei Menschen untersucht. Bei der Untersuchung werden zwei Experimente aus der Studie "Processing of expected values in human visual cortex" [5] von Dr. Philipp Kallerhoff reproduziert. Auf der Basis von deren Ergebnissen wird mithilfe von Hidden Markov Modellen nach möglicher Interpretation des Verhaltens gesucht.

1. Einleitung

Ein unerlässlicher Teil des menschlichen Lebens besteht darin Entscheidungen zu treffen. Daraus entstehen bestimmte Verhaltensmuster, die von Menschen unbewusst für weitere ähnliche Situationen wieder verwendet werden. Interessant wäre eine rationale Erklärung zu finden, wie diese Strategien entstehen und von welchen Faktoren sie beeinflusst werden.

Deswegen ist dieses Projekt, das einen Beitrag zur wissenschaftlichen Untersuchung von Entscheidungsprozessen bei den Menschen leistet, von Interesse für die gegenwärtigen Forschungsdiskussion auf diesem Gebiet. Die Untersuchung basiert auf zwei Experimenten, die auf einer Studie von Sharma et al [4] aufbauen.

In einer früheren Studie zeigte Sharma et al. (2003) [4], dass die neuronale Aktivierung im primären visuellen Cortex von Affen durch die Wahrscheinlichkeit der Belohnung eines gewählten Zieles moduliert wird. 2006 belegte Shuler et al. in einer anderen Studie, dass die Aktivierung des primären visuellen Cortex bei Raten ebenfalls durch die Belohnung moduliert wird.

Diese Studien bestätigen die Annahme, dass das Treffen von Entscheidungen abhängig von der Belohnung und deren Wahrscheinlichkeit ist. Auf der Basis dieser Annahmen wird in dieser Arbeit mit Hilfe ähnlicher Experimente versucht, das menschliche Verhalten zu analysieren und zu approximieren.

Im Folgenden werden die Versuche beschrieben, die Ergebnisse dargestellt und diskutiert. Anschließend wird auf den Einsatz des Hidden Markov Modells näher eingegangen. Dabei wird anhand der gewonnenen Daten ein Hidden Markov Modell gesucht, dass das Verhalten der Versuchspersonen möglichst gut approximiert.

2. Tools und Methoden

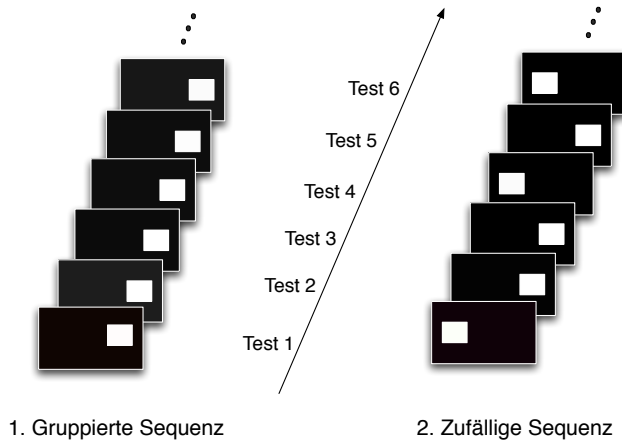
Die Experimente wurden mithilfe des "ViSaGe" - Systems [9] durchgeführt. Dabei handelt es sich um ein programmierbares Hardware- und Softwaresystem für PC-Systeme. Die Besonderheit dieses Geräts liegt darin, die kalibrierten visuellen Stimuli sowohl mit Präzision darzustellen als auch mit anderen externen Geräten zu synchronisieren. Die Reaktionen der Versuchspersonen werden mithilfe eines kabellosen Buttonbox aufgezeichnet.

An den Experimenten nahmen 8 Studenten im Alter zwischen 23 und 27 Jahre teil. Deren Aufgabe war es, vorherzusagen auf welcher Seite des Bildschirms Test-Stimuli erscheinen werden.

Die Durchführung der Experimente bei jedem Teilnehmer lässt sich in drei Phasen zusammenfassen. Zuerst wurden Erklärungen über den Ablauf des Experiments gegeben. Dann wurde ein kurzer Test, der sich dem Experiment ähnelt vom Teilnehmer absolviert, um sie für ihre Aufgabe zu trainieren. Schließlich wur-

de das entsprechende Experiment ohne Unterbrechung durchgeführt.

Insgesamt wurden zwei Experimente pro Teilnehmer hintereinander durchgeführt. Obwohl die Experimente vom Ablauf identisch sind, unterscheiden sie sich in der Form des Feedbacks nach jedem Vorhersageversuch.



1. Gruppierte Sequenz

2. Zufällige Sequenz

Abbildung 1. Beispielversuche mit 2 möglichen Stimuluslagen. 1. bezeichnet eine gruppierte Testsequenz mit Stimuli auf der rechten Seite. 2. ist eine Testsequenz mit zufälligen Stimuluslagen. Ein Stimulus bezeichnet ein 25x25 Pixel weißes Viereck auf einem schwarzen Bildschirm

Ein Versuch besteht aus einer Reihe von vierzig Blöcken. Die Blöcke umfassen Sequenzen von zufälligen oder gruppierten Stimuli in einer zufälligen Reihenfolge zufälliger Länge zwischen 5 und 10 (Siehe: Abbildung 2). Bei den einzelnen Test-Stimuli wird auf die Antwort vom Benutzer gewartet, der über das Betätigen der Fernbedienung eine Vorhersage darüber trifft, wo das nächste Stimuli am Bildschirm erscheinen wird.

Die Durchführung der Experimente lässt sich in 2 Teile untergliedern. Im ersten Teil (Experiment 1) bekommen die Teilnehmer ein positives Feedback in Textform: "Richtig" für eine zutreffende Vorhersage, "Falsch" sonst. Im zweiten Teil (Experiment 2) werden nur richtige Vorhersagen bei einer bestimmten Stimulusposition mit einem Piepton belohnt. Wobei die-

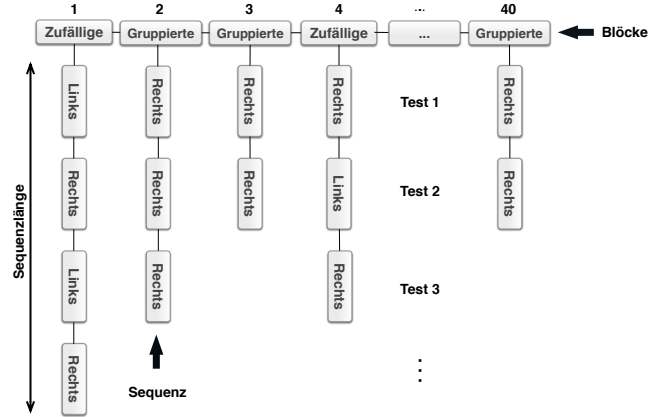


Abbildung 2. Interner Aufbau eines Versuches

se Stimulusposition im Laufe der einzelnen Versuchen unverändert bleibt.

Ein Hidden Markov Modell (HMM) ist ein Modell für eine gemeinsame Verteilung einer Menge von Zufallsvariablen $O_1, \dots, O_T, X_1, \dots, X_T$. Die O_i sind entweder stetig oder diskret verteilt und sichtbar; die X_i sind diskret verteilt und unsichtbar [11]. Dieses Modell fasst sich in einem 5-Tupel $\lambda = \langle S, K, \pi, A, B \rangle$ zusammen. Dabei ist

- $S = s_1, \dots, s_N$: die Menge der Zustände
- K : das Ausgabealphabet
- π : die Menge der Anfangswahrscheinlichkeiten für die Startzustände. Die Wahrscheinlichkeiten, dass sich das System zum Startzeitpunkt in den jeweiligen Zuständen befindet, werden mit $\pi = (\pi_1, \dots, \pi_N)$ bezeichnet und müssen folgende Bedingung erfüllen:

$$\sum_{i=1}^N \pi_i = 1. \quad (1)$$

- A : repräsentiert die Wahrscheinlichkeiten aller möglichen Zustandsübergänge in Form einer $N \times N$ -Matrix $A = (a_{ij})$. a_{ij} bezeichnet die Wahrscheinlichkeit, vom Zustand i in den Zustand j überzugehen. Es gilt die Nebenbedingung:

$$\sum_{j=1}^N a_{ij} = 1, \quad 1 \leq i \leq N. \quad (2)$$

- B: die Menge der Wahrscheinlichkeiten für die Ausgabe eines bestimmten Symbols $k \in K$ in einem Zustand s_i ; $b_i(k) = P(o_t = k | X_t = S_i)$.

In der Anwendung von Hidden Markov Modellen zur Analyse des Verhaltens der Versuchspersonen sollen mit Hilfe der Beobachtungen die Parameter eines passenden Hidden Markov Modells (θ) gefunden werden, die die Wahrscheinlichkeit $P((O_1, \dots, O_T) | \theta)$ maximieren.

Einen möglichen Lösungsweg bietet der Baum-Welch Algorithmus. Dieser Algorithmus berechnet iterativ einen Maximum-Likelihood Schätzwert für ein Modell θ eines HMM bei einer gegebenen Realisierung \mathcal{O} :

$$\arg \max_{\theta} P(\mathcal{O} | \theta) \quad (3)$$

Da die Wahrscheinlichkeitsdichte $P(\mathcal{O} | \theta)$ so klein werden kann, dass die Genauigkeit der Rechnerdarstellung nicht mehr ausreichend wird, wurde der Logarithmus der Wahrscheinlichkeitsdichte als Klassifizierungsfunktion gewählt. Somit lässt sich die Klassifikationsaufgabe aufgrund der monoton wachsenden Eigenschaft der logarithmischen Funktion als das Minimieren des Logarithmus der Wahrscheinlichkeitsdichte auffassen.

Der Baum-Welch Algorithmus dient dazu, ein Hidden Markov Modell zu trainieren, was bedeutet, die Parameter des HMM so zu setzen, dass es ein bestimmtes Phänomen so gut wie möglich modelliert. Der Algorithmus funktioniert so, dass basierend auf einem beliebigen anfänglichen Modell die Wahrscheinlichkeit der Realisation berechnet wird. Während der Berechnung wird festgehalten, wie oft Übergänge und Ausgabesymbole verwendet wurden. Im nächsten Schritt werden die Parameter so neu berechnet, dass häufiger benutzte Übergänge und Ausgabesymbole höhere Wahrscheinlichkeiten erhalten als weniger häufig benutzte. Nach der Anpassung wird das Modell für die Realisierung eine höhere Wahrscheinlichkeit errechnen, also besser angepasst sein. Die Schritte werden so oft wiederholt, bis sich die Modellparameter nur noch unwesentlich ändern [11].

Für die Umsetzung des Baum-Welch Algorithmus wurde die GHMM [3] Bibliothek eingesetzt. Es handelt sich um eine Open Source Bibliothek, die effiziente Datenstrukturen und Algorithmen für HMM mit diskreten und kontinuierlichen Emissionen umfasst.

Zuerst sollte die Topologie des HMMs festgelegt werden. Ein Modell mit viel mehr Zuständen als nötig benötigt mehr Trainingsdaten und entsprechend viel Ressourcen für die Abschätzung der HMM-Parameter. Im wiederum ein Modell mit nicht genügenden Zuständen erlaubt nicht die Erfassung einer akkuraten Wahrscheinlichkeitsverteilung der Trainingsdaten durch das HMM [2]. Wir testen pro Experiment zwei Konfigurationen, einen mit 2 und eine mit 3 hidden states. Die observierten Zustände sind 'L' für Links und 'R' für Rechts. Abbildung 3 stellt die ausgewählte Topologie dar. Zunächst wurde für je-

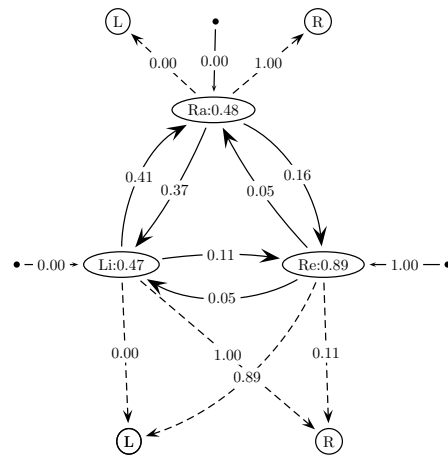


Abbildung 3. Beispiel eines HMMs: Die gestrichelten Linien repräsentieren die Emissionen mit den entsprechenden Wahrscheinlichkeiten. Die Ellipsen stellen die verborgenen Zustände dar: Li (Links), Ra (Random), Re (Rechts). Die eingekreisten 'R' und 'L' bilden die Ausgabewerte. Die Anfangswahrscheinlichkeiten sind auf den Zuständen eingehenden Pfeile ausgedrückt. Die Wahrscheinlichkeiten der Zustandsübergänge sind auf den durchgehenden Linien zwischen den Zuständen zu entnehmen.

der Versuchsperson ein Modell mit zufälligen Parametern generiert. Das Modell wurde mit dem 2/3 Anteil der zugehörigen beobachteten Sequenz mithilfe des Baum-Welch Algorithmus trainiert. Mit dem gelern-

ten Modell wurde zuerst die Log-Wahrscheinlichkeit (S2) der restlichen 1/3 der Soll-Sequenz ermittelt. Danach wurde die Log-Wahrscheinlichkeit des jeweils letzten 1/3 Anteils der von den Versuchspersonen generierten sowie entsprechenden beobachteten Sequenzen ermittelt. Zuletzt wurden 100 Sequenzen mit Hilfe des trainierten Modells generiert und deren Log-Wahrscheinlichkeit berechnet.

3. Ergebnisse

3.1. Psychophysik

Bei den zufälligen Sequenzen bleiben die Vorhersagen fast konstant über die gesamte Sequenz und liegen zwischen 39% und 48%. Dabei variiert die Standardfehlerquote des arithmetischen Mittels (SAM) zwischen 0.56% und 3.68%. Die Vorhersagen bleiben auf zufallslevel.

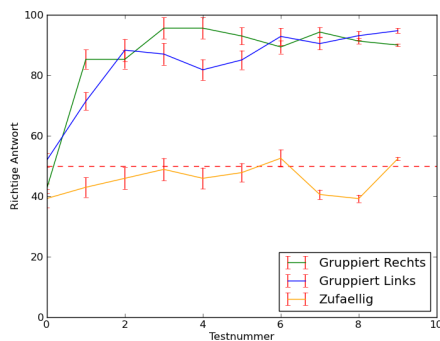


Abbildung 4. Experiment 1: Visuelle Auswertung der Testergebnisse. Die Fehlerbalken bezeichnen die Standardfehler der arithmetischen Mittels. Die arithmetischen Werte sind in Prozent dargestellt.

Bei Experiment 1 liegen die Vorhersagen bei den rechts gruppierten Sequenzen zwischen 41% und maximal 97% mit einem Wechsel der SAM in dem Bereich 0.4% bis 3.5%. Analog dazu liegen die Vorhersagen bei den links gruppierten Sequenzen in dem Bereich 50% bis 97%. Die SAM-Werte liegen zwischen 0.5% bis 3.7%. Es deutet auf eine Steigerung der Performance.

Die rechts gruppierten Vorhersagen bei Experiment

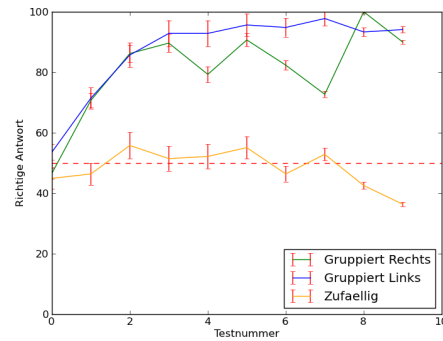


Abbildung 5. Experiment 2: Visuelle Auswertung der Testergebnisse. Die Fehlerbalken bezeichnen die Standardfehler des arithmetischen Mittels. Die arithmetischen Werte sind in Prozent dargestellt.

2 schwanken zu sehr und erreichen einen Durchschnitt von 99.8%. Daneben sind die links gruppierten Vorhersagen am Anfang monoton wachsend und verlaufen später asymptotisch zu einem maximalen Wert von ca. 98%. Dabei liegen die SAM-Werte zwischen 0.5% und 3.9%. Außerdem verlaufen die zufälligen Vorhersagen fast konstant und fallen zu einem minimalen Wert von 32%.

3.2. Modelle

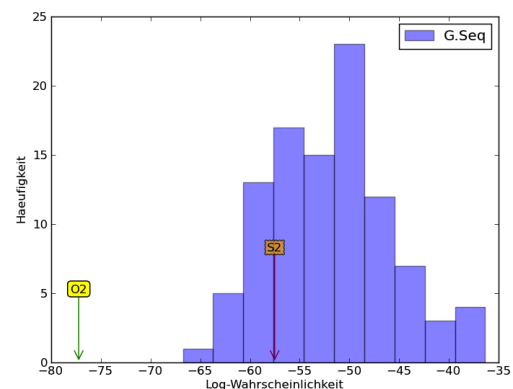


Abbildung 6. Experiment 1: Darstellung der ermittelten Log-Wahrscheinlichkeiten für Modell 4 (HMM mit 2 hidden states)

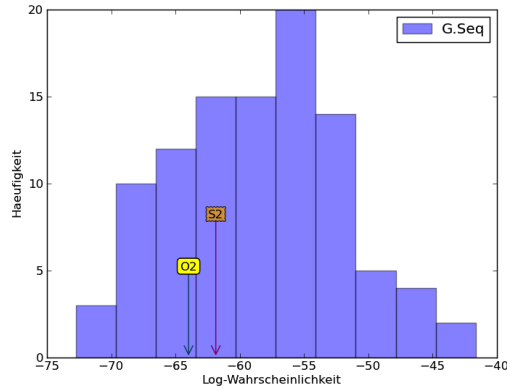


Abbildung 7. Experiment 1: Darstellung der ermittelten Log-Wahrscheinlichkeiten für Modell 7. HMM mit 3 hidden states.

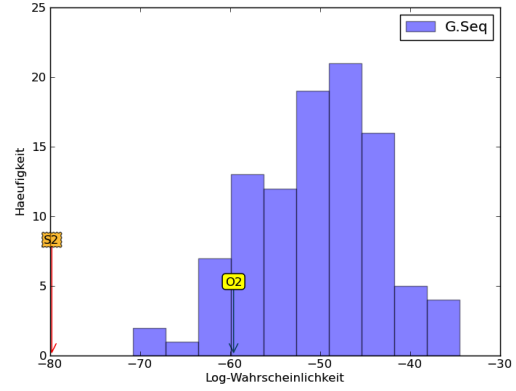


Abbildung 9. Experiment 2: Darstellung der ermittelten Log-Wahrscheinlichkeiten für Modell 2. HMM mit 3 hidden states.

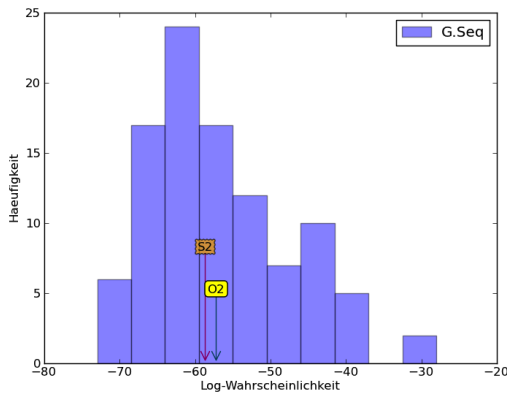


Abbildung 8. Experiment 2: Darstellung der ermittelten Log-Wahrscheinlichkeiten für Modell 3. HMM mit 3 hidden states.

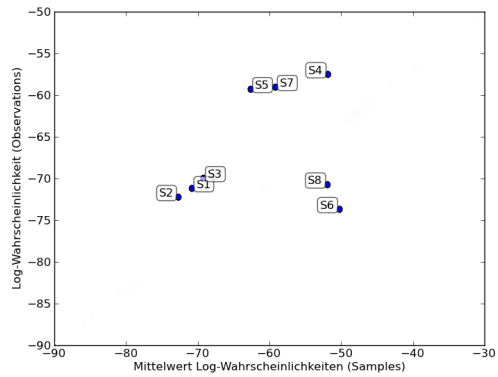


Abbildung 10. Experiment 1: Hidden Markov Modell mit 2 hidden states

Abbildungen (6 bis 9) stellen die Ergebnisse der Evaluierung der unterschiedlich generierten Modelle dar. Die blauen Balken widerspiegeln die Log-Wahrscheinlichkeit der vom trainierten Modell generierten Sequenz (G.Seq). Für ein Modell λ_i und eine von der Versuchsperson i generierte Sequenz O , O_i repräsentiert die Log-Wahrscheinlichkeit $\log(P(O|\lambda_i))$. S_i steht für die Log-Wahrscheinlichkeit $\log(P(S|\lambda_i))$ einer observierten Sequenz S bei der Versuchsperson i . Anhand dieser Evaluation werden die Sequenzen bezüglich der Modelle auf Kompatibilität untersucht. Bei dieser Klassifikationsaufgabe werden die ermittelten log-

Wahrscheinlichkeiten mit den vom Modell verglichen.

Aus den ermittelten Log-Wahrscheinlichkeiten der 100 getesteten Sequenzen wurde pro Modell je ein Mittelwert M_i ermittelt. Die Punkte $S_i(M_i, S2_i)$ auf Abbildungen 10 bis 13 beschreiben die Abhängigkeit von dem Log-Wahrscheinlichkeitswert der observierten Sequenzen ($S2_i$) zum jeweiligen Mittelwert (M_i). Aus diesen Abbildungen werden die Kandidaten zu den entsprechenden idealen Modellen ausgewählt.

Die Ausgabewahrscheinlichkeiten bei den idealen Modellen entstehen aus der Aufrundung der Ausgabewahrscheinlichkeiten des jeweilig gewählten Mo-

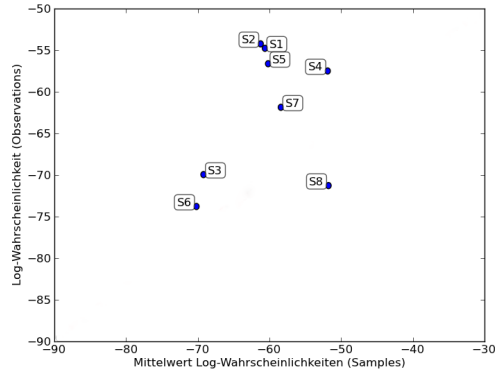


Abbildung 11. Experiment 1: Hidden Markov Modell mit 3 hidden states

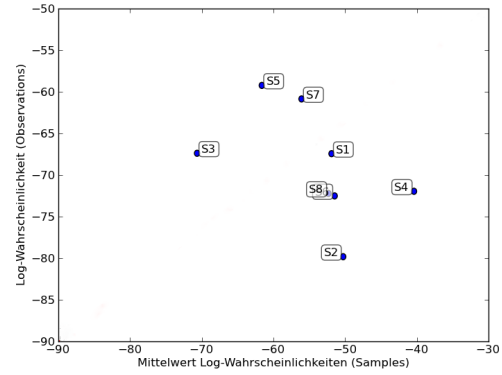


Abbildung 13. Experiment 2: Hidden Markov Modell mit 3 hidden states

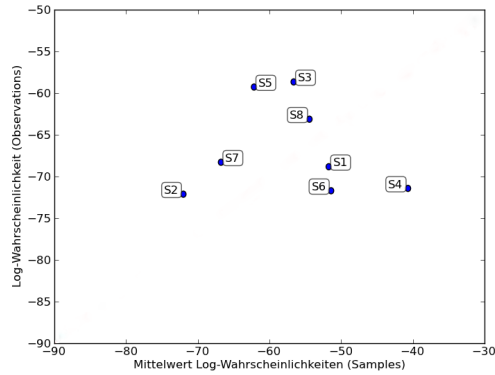


Abbildung 12. Experiment 2: Hidden Markov Modell mit 2 hidden states

dells θ_i mit dem passenden S_i aus Abbildungen 10 bis 13. Die Übergangswahrscheinlichkeiten sowie die Anfangswahrscheinlichkeiten entsprechen den von ausgewählten Modellen. Die Hidden States werden nach der Verteilung der Ausgabewahrscheinlichkeiten benannt. Die Rechts und Links hidden states wiederpiegeln die gruppierten Sequenzen und die zufälligen Sequenzen werden durch den hidden state Random abgebildet.

4. Diskussion

Insgesamt entsprechen die Ergebnisse den Erwartungen. Ähnliche Ergebnisse wurden von Dr.

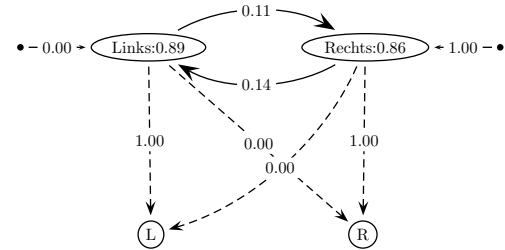


Abbildung 14. Experiment 1: Ideales Modell mit 2 hidden states

Kallerhoff [5] erfasst. Bei den gruppierten Sequenzen wachsen die richtigen Vorhersagen im Laufe des Blockes. Mit zunehmender Wiederholung steigt die Chance, dass die Versuchspersonen die richtige Seite treffen. Die Schwankungen bei den gruppierten Sequenzen in Experiment 2 sind verursacht durch die heterogene Verteilung der Samples bei der gruppierten Sequenzen. In der 9^{ten} Versuchsreihe zum Beispiel sind viel weniger Test-Stimuli als in den anderen Reihen generiert worden. Bei den zufälligen Sequenzen ist eine Steigerung der Performance nicht zu erwarten. Dabei wechselt die Trefferquote bei beiden Experimenten hin und her um 50% und erreicht 30% bei Experiment 2. Die Ergebnisse stimmen qualitativ mit dem Bayes'schen Modell von Dr. Kallerhoff [5, 35].

Für die unterschiedlich vom trainierten Modell

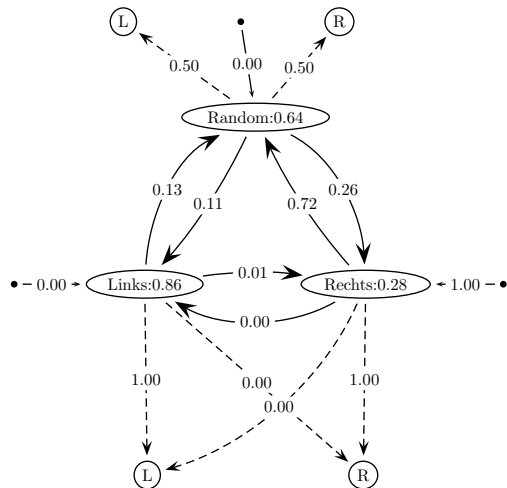


Abbildung 15. Experiment 1: Ideales Modell mit 3 hidden states

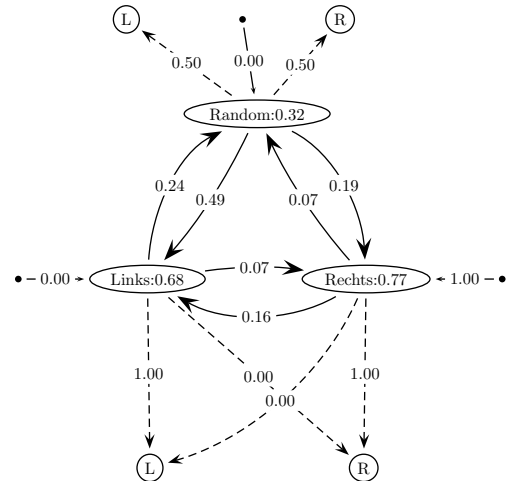


Abbildung 17. Experiment 2: Ideales Modell mit 3 hidden states

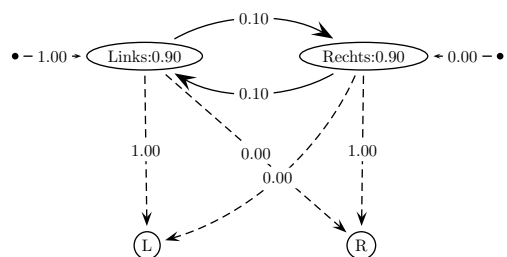


Abbildung 16. Experiment 2: Ideales Modell mit 2 hidden states

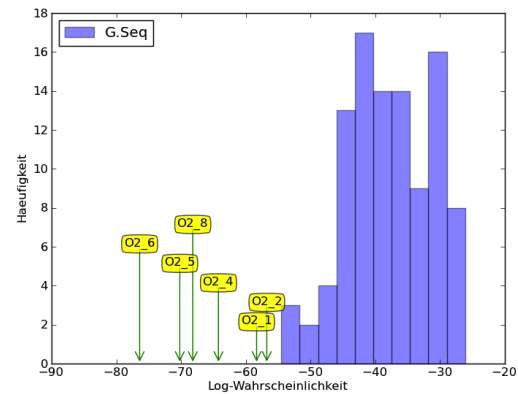


Abbildung 18. Experiment 1: Darstellung der ermittelten Log-Likelihoods für die unterschiedlichen Subjects O_i mithilfe des idealen Hidden Markov Modells mit 2 hidden states

generierten Sequenzen fällt auf, dass die Ergebnisse stark unterschiedlich zueinander sind. Es könnte darauf zurückzuführen haben, dass der Baum-Welch-Algorithmus in der Regel nur eine lokale optimale Lösung findet. Die Parameter, die der Algorithmus liefert, hängen also von der Wahl der Initialparameter ab [10].

Generell liegen alle Modelle den entsprechenden von den Versuchsperson generierten Sequenzen zugrunde. Aufgrund dessen kleinster Log-Wahrscheinlichkeit wurde für Experiment 1 das Modell θ_4 ausgewählt. Die Modelle θ_3 und θ_7 bei Experiment 2 wurden jeweils für das Modell mit 2 bzw. 3 hidden states ausgewählt. Diese Modelle bilden eine gute Grundlage für die Beschreibung des

Verhaltens der Versuchspersonen.

Bei Experiment 1 lässt sich das Verhalten der Versuchspersonen zu 75% mit einem Modell mit 3 hidden states approximieren gegen 50% bei Experiment 2. Das Verhalten der Versuchsperson lässt sich nicht mit einem Hidden Markov Modell mit 2 hidden states approximieren. Offensichtlich neigen die Versuchspersonen bei beiden Experimenten dazu, mit der rechten Seite anzufangen. Besonders auffällig ist, dass bei Experiment 1 die Wahrscheinlichkeit

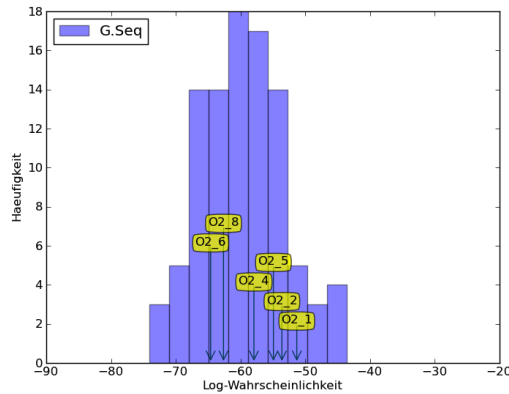


Abbildung 19. Experiment 1: Darstellung der ermittelten Log-Likelihoods für die unterschiedlichen Subjects O_i mithilfe des idealen Hidden Markov Modells mit 3 hidden states

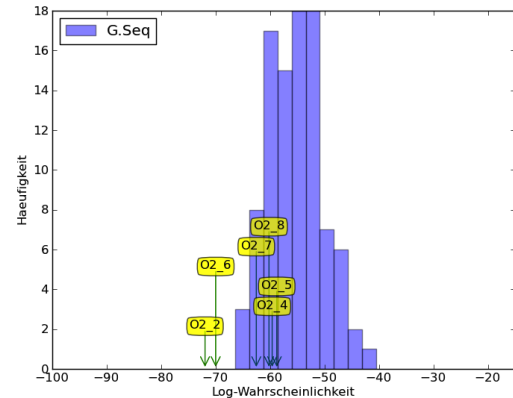


Abbildung 21. Experiment 2: Darstellung der ermittelten Log-Likelihoods für die unterschiedlichen Subjects O_i mithilfe des idealen Hidden Markov Modells mit 3 hidden states

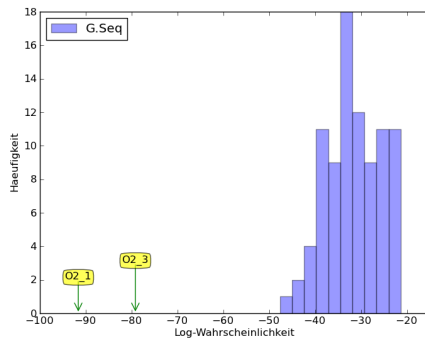


Abbildung 20. Experiment 2: Darstellung der ermittelten Log-Likelihoods für die unterschiedlichen Subjects O_i mithilfe des idealen Hidden Markov Modells mit 2 hidden states

eines Wechsels bei einem Treffer auf der linken Seite gering ist. Bei Experiment 2 ist die Wahrscheinlichkeit eines Wechsels bei den Treffern auf der linken oder auf der rechten Seite ebenfalls gering.

Man kann erschliessen, dass die Versuchspersonen bei den Experimenten eine einfache Strategie benutzt haben, die sich stark nach der Belohnung orientiert.

5. Referenzen

Literatur

- [1] M. Bicego and V. Murino. Investigating hidden markov models' capabilities in 2d shape classification. *IEEE Transactions on pattern analysis and machine intelligence*, 26(2), February 2004.
- [2] R. J. Elliot, L. Aggoun, and J. Moore. *Hidden Markov Models, Estimation and Control*. Springer-Verlag, New York, 1995.
- [3] ghmm.org. <http://www.ghmm.org>.
- [4] S. J. D. V, T. JB, M. EK, and S. M. *V1 neurons signal acquisition of an internal representation of stimulus location. V1 neurons signal acquisition of an internal representation of stimulus location*. Department of Brain and Cognitive Sciences, 2003.
- [5] P. Kallerhoff. *Processing of expected values in human visual cortex*. PhD thesis, Technische Universität Berlin, June 2009.
- [6] D. N. M. R. P.V. Bork, H. Grote. *Data Analysis Techniques in High Energy Physics Experiments*. Cambridge University Press, 1993.
- [7] L. R. Rabiner and Fellow. *A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition*. IEEE.
- [8] J. R. Taylor. *An Introduction to Error Analysis: The Study of Uncertainties if Physical Measurements*. University Science Books, 1982.
- [9] C. University. <http://www.crs ltd.com/catalog/visage>, August 2011.

- [10] B. Wichern. *Hidden-Markov-Modelle zur Analyse und Simulation von Finanzzeitreihen*. PhD thesis, Universität zu Köln, 2001.
- [11] H. Wunsch. Der baum-welch algorithmus für hidden markov models, ein spezialfall des em-algorithmus.