



GEORG-AUGUST-UNIVERSITÄT  
GÖTTINGEN

Hausarbeit zur Veranstaltung  
*Proseminar Artificial Life*  
(Veranstaltungsnummer 990195)  
im Sommersemester 2014

Thema:

**Ausbalancieren von Kosten und Nutzen von  
Lernfähigkeit**

auf Grundlage der Publikation von Kai Olav Ellefsen zur ECAL 2013

**Balancing the Costs and Benefits of Learning Ability**

eingereicht von:

**Steffen Brandl**

Matrikelnr: 21162403

Prüfer:

**Prof. Dr. Winfried Kurth**

Datum:

**9. November 2014**

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Hintergrund</b>	<b>1</b>
2.1	Lernen - Kosten und Nutzen . . . . .	1
2.2	Modellierung der Kosten von Plastizität . . . . .	2
2.3	Der Baldwin Effekt - Regulierung der Plastizität über Generationen . . .	3
2.4	Rechnerische Modellierung des Baldwin Effekts . . . . .	4
2.5	Sensitive Phasen - Regulierung von Plastizität innerhalb eines Individuums	5
2.6	Rechnerische Modellierung von sensitiven Phasen . . . . .	5
<b>3</b>	<b>Hypothese</b>	<b>5</b>
<b>4</b>	<b>Versuchsaufbau</b>	<b>6</b>
4.1	Die Umgebung . . . . .	6
4.2	Die Agenten . . . . .	7
4.3	Plastizität . . . . .	8
4.4	Evolutions-Algorithmus . . . . .	8
<b>5</b>	<b>Ergebnisse</b>	<b>9</b>
5.1	Veränderungsraten und Lernleistung . . . . .	9
5.2	Statische oder dynamische Plastizität . . . . .	10
<b>6</b>	<b>Schlussfolgerung</b>	<b>12</b>
<b>7</b>	<b>Weiterführende Forschung</b>	<b>14</b>
	<b>Literaturverzeichnis</b>	<b>15</b>

# 1 Einleitung

Bereits seit mehreren Jahrhunderten beschäftigen sich Wissenschaftler mit der Evolution. Eine wichtige Rolle spielt dabei die Fähigkeit eines Individuums zum Lernen, da dadurch die jeweilige Fitness zunimmt. Beim Lernen entstehen jedoch Kosten, welche so weit wie möglich reduziert werden. Das Ziel ist, dass so viel wie nötig, aber so wenig wie möglich gelernt wird und somit die Plastizität durch angeborene Strategien ersetzt werden kann. Lernprozesse sollten also nur stattfinden, wenn sie dem Lebewesen den größten Nutzen bringen. Neuronale Plastizität bezeichnet die Fähigkeit von Synapsen und Nervenzellen, sich in ihren Eigenschaften zu verändern und somit neu gelernte Verhaltensmuster zu speichern und anzuwenden.

Diese Hausarbeit basiert auf einer Veröffentlichung von Kai Olav Ellefsen, welche er 2013 auf der zwölften ECAL (European Conference on Artificial Life) vorstellte. In dieser untersuchte er einerseits den *Baldwin Effekt*, welcher die Plastizität über Generationen reguliert, und andererseits die *sensitiven Phasen*, welche dieses innerhalb der Lebensdauer eines Individuums tun.

Diese beiden Regulierungsprozesse werden im Folgenden genauer erklärt und die zugehörige Hypothese von Ellefsen vorgestellt. Im Folgenden wird der Versuchsaufbau und die Durchführung erläutert und die Ergebnisse vorgestellt.

## 2 Hintergrund

### 2.1 Lernen - Kosten und Nutzen

Die Vorteile des Lernens wurden bereits in vielen vorangegangenen Studien über die Interaktion zwischen Evolution und Lernen dokumentiert (siehe dazu [8, 14, 17]). Bei solchen Studien ist es allerdings entscheidend, auf die durch das Lernen entstandenen Kosten zu achten. Die endgültige Lernstrategie wird anhand des Verhältnisses von Kosten und Nutzen ausgewählt. Die darauffolgende Generation von Individuen ist durch einen evolutionären Prozess bereits den Veränderungen der Umwelt angepasst.

Einen Überblick über die Kosten und den Nutzen gibt Giles Mayley im Jahr 1996 [15]. Die Nutzen sind offensichtlich, wie zum Beispiel die Steigerung der Fitness. Als Kosten sind beispielsweise die energetischen Kosten, die beim Lernen aufgebracht werden müssen, und die erneut aufzubringende Energie bei Fehlversuchen, die bei unbekanntem Umwelteinflüssen auftreten können, zu nennen.

Eine Folgerung aus den Kosten für die Lernfähigkeit ist, dass diese einen adaptiven Wert für das Individuum haben müssen. Wenn es möglich ist, sollte die natürliche Selektion die Kosten reduzieren, indem plastisches Verhalten durch genetisch vererbte Mechanismen ersetzt wird.

Kosten und Nutzen variieren stark zwischen Individuen gleicher Art, so kann man beispielsweise den Unterschied beobachten, dass junge Individuen einen größeren Lernnutzen bei Veränderungen haben als erwachsene, da sich diese bereits alle wichtigen Eigenschaften und Verhaltensmuster angeeignet haben und diese schwerer wieder geändert werden

können. Selbst innerhalb eines Individuums variiert das Kosten/Nutzen-Verhältnis in verschiedenen Situationen. Bei wichtigen Veränderung sind die Nutzen deutlich höher als bei unwichtigen Situationen.

Peter Turney verglich 1996 zehn Eigenschaften von phänotypischer Plastizität (Lernaktivitäten) und phänotypischer Unnachgiebigkeit (Instinkthandlungen) [20]. Zwei dieser Eigenschaften sind die Länge der Lernphasen und die Varianz/Zuverlässigkeit. Handlungen aus Instinkt weisen kurze Lernphasen auf und sind invariant, weisen aber eine hohe Zuverlässigkeit auf. Plastizität hingegen hat längere Lernphasen und ist variant, dafür aber unzuverlässiger.

## 2.2 Modellierung der Kosten von Plastizität

Benjamin Kerr und Marcus Feldmann untersuchten, wie sich die Zuverlässigkeit von Stimuli auf die Brauchbarkeit des Langzeitgedächtnisses auswirkt [11]. Die beiden argumentierten, dass die Variabilität der Umgebung der bestimmende Faktor für die Entscheidung zum evolutionären Vorteil des Lernens ist. Sie behaupteten, dass die Beziehung zwischen der Veränderlichkeit der Umgebung und die Nützlichkeit des Lernens dem *Goldlöckchen-Prinzip* folgt. Dieses besagt, die Veränderlichkeit der Umgebung müsse sich in einem perfekten Zustand befinden, damit sich lernen lohnen würde. Sie dürfe also nicht zu schwach oder zu stark ausgeprägt sein. Kerr und Feldman schlossen aus ihren Ergebnissen, dass bei dauerhaft schnellen Veränderungen der Nutzen des Kurzzeitgedächtnisses vorteilhaft ist, während bei einer stabileren Umgebung das Erinnerungsvermögen stärker genutzt wird. Sie mutmaßten sogar, dass in einer komplett zuverlässigen Welt, in der sich die Voraussetzungen niemals verändern, auch die Notwendigkeit des Lernens verloren gehen würde, da alle Verhaltensweisen bereits hart-codiert sein könnten. Sie bezogen ihre Forschung allerdings nicht auf mehrere Generationen, sodass die Möglichkeit von genetisch optimierten Verhaltensmustern nicht berücksichtigt wurde.

Aimee Dunlap und David Stephens lieferten 2009 die ersten experimentellen Ergebnisse bei einem Versuch über mehrere Generationen [5]. Sie nutzten dafür Populationen von Fruchtfliegen (*Drosophila melanogaster*) und untersuchten deren Verhalten über 30 Generationen. Dabei wiesen sie nach, dass die Fliegen bei einigen Umgebungsveränderungen mit Lerneffekten reagierten, während sie sich bei anderen Veränderungen dagegen entschieden. Durch ein Aversions-Lernexperiment wurden zwei Arten von Umgebungsveränderungen erkannt, welche den entwickelten Lerngrad der Fruchtfliegen unterschiedlich beeinflussten. Die erste Art nannten Dunlap und Stephens *best-action fixity*. Dieser Wert beschreibt, zu welchem Grad die beste durchführbare Aktion in der Umgebung immer die gleiche Aktion ist. Ein hoher Wert dieses Parameters weist also darauf hin, dass die Strategie, immer die gleiche Aktion durchzuführen, in den meisten Fällen zum Erfolg führt. Die zweite Art der Veränderung wurde als *reliability of experience* bezeichnet. Dieser Parameter beschreibt die Beständigkeit der Beziehung zwischen Erfahrung und bester Aktion. Der Wert gibt an, bis zu welchem Grad es möglich ist, assoziativ aversiv zu lernen. Dunlap und Stephens stellten in ihrem Experiment fest, dass sich in einer Situation mit hoher Verlässlichkeit der Erfahrung und geringer *best-action fixity* die Fliegen am häufigsten für einen Lernprozess entschieden. Bei der gegenteiligen Situation

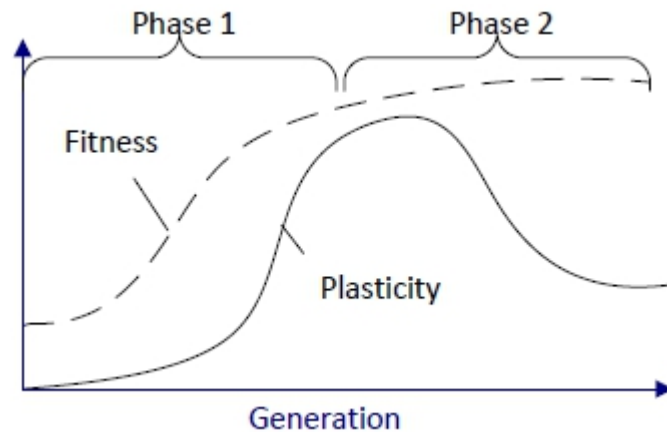


Abbildung 1: Die zwei Phasen des Baldwin Effekts. Quelle: [6]

wurden nicht-lernende fixe Strategien angewendet.

Kai Olav Ellefsen legte 2013 in seinen Experimenten, auf die später genauer eingegangen wird, den Fokus auf die Regulierung der *best-action fixity* [6]. Die beste Aktion, die einem Lebewesen die größte Fitness in einer Umgebung bringt, sollte verändert werden. Dabei ist es wichtig, dass das Feedback, welches dem Agenten mitteilt, ob die Aktion gut oder schlecht war, immer richtig ist. Er führte seine Experimente durch, da er glaubte, dass Dunlap und Stephens [5] einen Teil der Wirklichkeit übersehen hatten. Ellefsen behauptete, dass sich bei einem ausreichenden Abfall der Beständigkeit der besten Aktion gegen das Lernen entschieden wird. Dieses würde die Verwendbarkeit des *Goldlöckchen-Prinzips* bestätigen, welches Kerr und Feldman in die Forschung des Lernverhaltens übertrugen [11].

## 2.3 Der Baldwin Effekt - Regulierung der Plastizität über Generationen

Der Baldwin Effekt, benannt nach James Mark Baldwin (siehe dazu [2]), ist ein interessantes Beispiel dafür, wie die Evolution zur Kosteneinsparung die Plastizität über mehrere Generationen reguliert. Der Effekt weist darauf hin, wie gelernte Eigenschaften und Verhaltensweisen in das Erbgut von Individuen durch einen indirekten Mechanismus übernommen werden.

Der Baldwin Effekt besteht aus zwei Phasen (siehe dazu auch Abbildung 1). Die Voraussetzung für das Auftreten dieses Effekts ist eine Veränderung der Umgebung, welche die Population dazu zwingt, sich anzupassen. Während der ersten Phase wird die Evolutionsrate durch lernen beschleunigt. Dieses lässt sich dadurch begründen, dass durch Lerneffekte die schwächeren Individuen stärker werden, wodurch die Fitness-Landschaft geebnet wird. Geoffrey Hinton und Steven Nowlan führten dazu einen 1987 veröffentlichten Versuch durch, bei dem es nur eine einzige richtige Lösung gab und kein Fitness-Gradient existierte, mit dem die Evolution gesteuert werden konnte, wenn kein Lerneffekt auftrat [9]. In dieser ersten Phase des Baldwin Effekts gleichen die Vorteile

von Plastizität ihre eigenen Kosten aus, wodurch eine zunehmende plastische Population entsteht. Dieses sieht man in der Abbildung 1, in der zum Ende der ersten Phase die Fitness beinahe ihr Maximum erreicht und ein sehr hoher Wert bei der Plastizität zu erkennen ist.

In der zweiten Phase tritt die Auswirkung der genetischen Anpassung auf. Es werden also die gelernten Eigenschaften in das Erbgut übernommen, sodass diese nicht neu erlernt werden müssen. Dieses ist eine Folge der entstehenden Kosten des Lernens. Die Population ist nun voll mit Individuen, welche sich an die Veränderungen anpassen können und entwickelten Individuen, welchen die richtigen Eigenschaften bereits angeboren sind. Diese entwickelten Individuen haben nun den Vorteil, dass bei ihnen die Lernkosten nicht mehr entstehen. Durch diese Veränderungen innerhalb der Population sinkt ihre Plastizität. In einer weiteren Arbeit hob auch Giles Mayley (1996) hervor, dass es die variierende Kosten/Nutzen-Abwägung von Plastizität ist, die die Veränderungen des Lernlevels in einer sich entwickelnden Population erzwingt [16].

## 2.4 Rechnerische Modellierung des Baldwin Effekts

Es gab in der Vergangenheit einige Experimente zum Baldwin Effekt, bei denen einfache, sich entwickelnde Individuen verwendet wurden. Die meisten dieser Versuche wurden in einer sich nicht verändernden Umgebung durchgeführt, in die nicht angepasste Individuen eingesetzt wurden. Sobald Umgebungsveränderungen ermöglicht werden, wird die analytische Auswertung deutlich komplexer. Dennoch ist es nicht ausreichend, lediglich die Interaktion zwischen Lernen und Evolution in fixen Umgebungen zu untersuchen, wie es auch Russell Anderson (1995) bereits beschrieb [1]. In veränderbaren Umgebungen ist beispielsweise die Fähigkeit, Plastizität als Puffer gegen Veränderungen einzusetzen, deutlich leichter erkennbar.

Im Jahr 2002 untersuchte James Watson mit zwei weiteren Forschern die Beziehung zwischen der Komplexität und Stabilität einer Lernaufgabe und der Tendenz, dass genetische Anpassung auftritt [21]. Dabei fanden sie heraus, dass in sehr instabilen Umgebungen genetische Anpassung dem Lernen vorgezogen wird. Je geringer die Instabilität war, desto geringer war auch der Grad der genetischen Anpassung. Diese Feststellung erklärten sie so, dass in den Umgebungen mit den schnellsten Veränderungen die entwickelten Verhaltensmuster besser funktionierten, als die gelernten. In leicht stabileren Umgebungen stiegen die Vorteile vom Lernen an, sodass die Lernrate niemals komplett auf 0 sank.

Giles Mayley untersuchte 1996 die Effekte von zwei wichtigen Variablen und wie sie sich auf das Lernverhalten von Individuen über mehrere Generationen auswirken [15]. Zum einen sind die Kosten für das Lernen zu nennen und zum anderen die Korrelation zwischen Genotyp und Phänotyp. Er fand bei seiner Forschung heraus, dass sowohl eine *neighborhood* Korrelation zwischen Geno- und Phänotyp, als auch evolutionäre Kosten zum Lernen notwendig sind, damit eine genetische Anpassung stattfinden kann.

Takahiro Sasaki und Mario Tokoro veröffentlichten 1999 die Ergebnisse ihrer Forschungsarbeit über verschiedene Raten von Vererblichkeit von erworbenen Charakteristika und die Auswirkung von verschiedenen Veränderungsrate der Umgebung [18]. Dabei entdeckten sie in Umgebungen mit relativ kleinen Veränderungsrate Hinweise auf das Einwirken



Abbildung 2: Eine hypothetische Zuordnung von Lernstrategien zu verschiedenen Umgebungsstabilitätsraten. Quelle: [6]

des Baldwin Effekts. Bei Umgebungen mit relativ hohen Veränderungsraten hingegen fand keine genetische Anpassung statt. Daraus schlossen sie, dass der Baldwin Effekt einen gewissen Grad an Stabilität benötigt, um Wirkung zu zeigen. Diese Aussage wird durch die Experimente von Ellefsen gestützt [6].

## 2.5 Sensitive Phasen - Regulierung von Plastizität innerhalb eines Individuums

Eine sensitive Phase ist eine Phase im Leben eines Lebewesens, in der Umwelteinflüsse eine große Rolle bei der Entwicklung von vielen wichtigen Eigenschaften und Fähigkeiten einnehmen [13]. Ein sehr bekanntes Beispiel dafür lieferten David Hubel und Torsten Wiesel (1970) [10]. Sie nähten bei ihren Versuchen verschiedenen Katzen in verschiedenen Abschnitten ihres Lebens ein Auge zu und fanden dabei heraus, dass der Entzug der visuellen Reize in frühen Lebensstadien die Entwicklung des zugenähten Auges behinderte und diese Katzen auf dem Auge erblindeten.

## 2.6 Rechnerische Modellierung von sensitiven Phasen

John Bullinaria untersuchte 2003 sensitive Phasen als Teil einer Simulation des menschlichen okulomotorischen Systems [3]. Dabei erzeugte er durch einen Evolutions-Algorithmus ein Plastizitäts-Modell, welches die Lernfähigkeit in Abhängigkeit des Alters steuert. Diese altersabhängige Plastizität besitzt Parallelen mit biologischen sensitiven Phasen. Die interessanteste Eigenschaft von sensitiven Phasen ist die Lernkosten-Reduzierung durch Verkleinerung der plastischen Phasen eines Lebewesens (siehe auch [3, 12]). Ellefsen untersuchte in seinen Versuchen die Entwicklung von sensitiven Phasen in verschiedenen Kosten/Nutzen-Verhältnissen vom Lernen und verglich die genetischen Anpassungen in sensitiven Phasen mit denen in lebenslanger Plastizität [6].

## 3 Hypothese

Die Hypothese der Forschung von Ellefsen wird in Abbildung 2 dargestellt. Sie zeigt in einem Modell, wie der Baldwin Effekt, sensitive Phasen und das Kosten/Nutzen-Verhältnis

miteinander verknüpft sind. Durch die Regulierung der Rate der Umweltveränderung werden die Kosten und Nutzen des Lernens beeinflusst. Ellefsen stellt in Abbildung 2 seine Annahme dar, wie Lernstrategien verschiedenen Veränderungsraten zuzuordnen sind. In sich ständig verändernden Umgebungen hat das Lernen keinen Vorteil, da es keine langanhaltenden Regeln gibt, die zu erlernen sind. Ebenso lernen Lebewesen in komplett stabilen Umgebungen auch nichts, da es keine Veränderungen gibt.

Ellefsen nahm an, dass sich zwischen diesen beiden Extremen Individuen finden werden, welche verschiedene Plastizitätslevel aufweisen. Wenn zu einer stabilen Umgebung nur langsame Veränderungen hinzugefügt werden, können diese noch durch Evolution ausgeglichen werden. Wenn jedoch die Zeit zwischen Veränderungen auf ein bestimmtes Niveau abfällt, wird die Grenze erreicht, bis zu der genetische Anpassung vorteilhafter ist als zu lernen. Wenn die Veränderungen noch relativ gering sind, sollten sensitive Phasen im Leben eines Individuums ausreichend sein, sodass die Kosten für lebenslange Plastizität eingespart werden können. Wenn jedoch die Veränderungsrate relativ hoch ist, müssen sich die Individuen öfter anpassen und neue Verhaltensweisen lernen. Dafür ist eine lebenslange Lernfähigkeit unumgänglich. Sensitive Phasen können als Kompromiss zwischen der billigen und langsamen Veränderung des Genotyps und der teuren und schnellen Veränderung durch individuelles Lernen gesehen werden.

## 4 Versuchsaufbau

### 4.1 Die Umgebung

Ellefsen modifizierte bei seinem Versuch einen Versuchsaufbau von Peter Todd und Geoffrey Miller, um die Entwicklung von assoziativem Lernen zu untersuchen [19]. Bei dem Experiment wurde eine Unterwasserkreatur simuliert, welche ihr komplettes Leben an einer Stelle verbringt. An dieser Kreatur schwimmen zwei verschiedenfarbige Substanzen vorbei und es muss entschieden werden, ob die Substanz gefressen wird oder nicht. Eine der Substanzen ist giftig, die andere nahrhaft. Somit muss gelernt werden, welche der beiden Farben welche Eigenschaft besitzt. Die Assoziation zwischen Farbe und Genießbarkeit hängt von dem Futterbereich ab, in der eine Kreatur geboren wird. In verschiedenen Bereichen ist die Farbzuordnung umgekehrt, sodass nicht für alle Kreaturen eine einzige Strategie die beste ist.

Ellefsen untersuchte mit diesem Versuchsaufbau die Beziehung zwischen Plastizität und verschiedenen Umgebungsveränderungsraten. Dafür mussten zwei wichtige Erweiterungen durchgeführt werden. Erstens wurde das Lernen mit Kosten verbunden, die die biologischen Kosten darstellen, und zweitens verändert sich die Umgebung in festgelegten Intervallen, wobei dann die Farben der zwei Substanzen umgekehrt werden. Somit entsteht ein Unterschied zwischen den beiden Arten der Assoziationsänderung. Während im Originalversuch die Assoziation entsprechend des Futterbereichs variierte, variiert diese im zweiten Versuch über die Zeit. Die Rate, in der sich die Umgebung veränderte, wurde über eine einzige Variable gesteuert, die Stabilitätsperiode. Über diese wird bestimmt, wie viele Generationen zwischen zwei Veränderungen liegen. Werte unter 1 bedeuten,



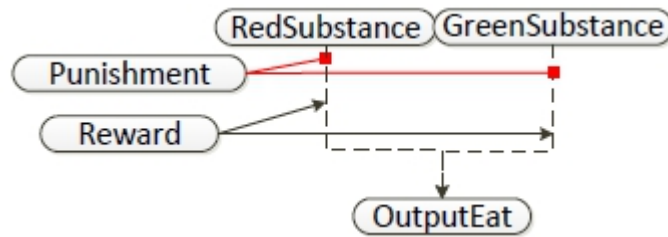


Abbildung 3: Das im Experiment verwendete neuronale Netz. Quelle: [6]

dass mehrere Veränderungen pro Generation auftreten, so sind beispielsweise bei einem Wert von 0,2 fünf Wechsel zu beobachten.

## 4.2 Die Agenten

Wie bereits Mayley feststellte, sind zwei Bedingungen in Experimenten über genetische Anpassung notwendig [16]. Erstens muss die Plastizität der Agenten unter genetischer Kontrolle stehen und zweitens müssen die durch Plastizität erzeugten Charakteristiken auch genetisch erzeugbar sein.

In Ellefsens Experiment sind diese beiden Bedingungen durch ein künstliches neuronales Netz gegeben. Dieses ist fähig sowohl festcodierte Regeln als auch neuromodulatorisches Lernen zu verwenden, wenn entschieden werden muss, welche Substanzen gefressen werden soll und welche nicht. Durch einen Evolutions-Algorithmus werden die Anfangswerte für die Verbindungsgewichte und die Lernraten festgelegt.

Das neuronale Netz ist in Abbildung 3 dargestellt. Die gestrichelten Verbindungen sind die plastischen Verbindungen, während die anderen Verbindungen für das gesamte Experiment festgelegt sind. Die Verbindungen, die anderen Verbindungen zugewiesen sind, sind die Neuromodulatoren. Diese beeinflussen die Lernraten in den dazugehörigen Verbindungen. Durch diesen Aufbau wird durch Wahrnehmung von Belohnungen bzw. Bestrafungen *reinforcement learning* erreicht. Sobald sich in der Umgebung die Zuordnung der Farben zu den Substanzen ändert, bemerken die Agenten, dass ihr Fressverhalten zu einem anderen Feedback führt als bisher, und ändern ihre Präferenzen entsprechend ihrer neuromodulatorischen Plastizität.

Verbindungen im neuronalen Netz werden mit folgender Formel aktualisiert:

$$\Delta w_{i,j} = \eta * mod * |x_i x_j|$$

In der Gleichung stellen  $\eta$  die aktuelle Lernrate und  $mod$  die Stärke der ankommenden Neuromodulation dar.  $x_i x_j$  ist das Produkt der pre- und post-synaptischen Aktivität, also ein regulärer Hebbian-Term.

Der absolute Wert des Hebbian-Update-Terms wird für die Berechnung des neuen Gewichts verwendet, damit nur das modulatorische Signal entscheidet, in welche Richtung sich der Wert verändert. Bei negativer Modulation sinkt das Gewicht der Verbindung und bei positiver steigt das Gewicht entsprechend. Sollte keine Modulation vorliegen, wird der Wert nicht verändert.

Parameter	Value
Generations	200
Adults	15
Children	25
Crossover probability	0
Mutation probability	0.01
Genes per individual	3 (static) or 60 (dynamic)
Bits per Gene	8
Elite fraction	0.1
Culling fraction	0.1

Tabelle 1: Die genutzten Parameter für den Evolutions-Algorithmus. Quelle: [6]

### 4.3 Plastizität

Die Experimente wurden mit zwei verschiedenen Plastizitäts-Typen durchgeführt. Der erste Typ war die *statische Plastizität*, bei der die Plastizität über die gesamte Lebensdauer gleich war, und über eine einzige Lernrate gesteuert wurde. Für den zweiten Typen, die *dynamische Plastizität*, wurde eine Funktion entwickelt, welche den Plastizitätslevel eines Agenten über sein Leben in einem Zwei-Zeitschritt-Intervall regelt. Das bedeutet, dass in dem Leben eines Agenten, welches aus 100 Zeitschritten besteht, das Plastizitätslevel 50 Mal geändert werden kann. Damit eine gleichmäßige, altersabhängige Funktion erstellt werden kann, werden für jeden Wert die nächsten sieben Zeitschritte beobachtet. Jedem Zeitschritt wird ein Wert zwischen -2 und 2 zugeordnet und aus den acht Zeitschritten dann der Durchschnittswert der Plastizität berechnet.

### 4.4 Evolutions-Algorithmus

Es wurde das System SEVANN (Script-Based Evolution of Artificial Neural Networks) für die Entwicklung der Lernraten verwendet. Es handelt sich dabei um ein System, mit dem Experimente zu der Evolution von Parametern und Topologien in neuronalen Netzen genutzt werden können. Es wurde 2010 von Ellefsen und Keith Downing vorgestellt (siehe dazu [4]). Die für das Experiment genutzten Parameter sind in Tabelle 1 dargestellt.

Die Ergebnisse waren am stabilsten, wenn *crossover* ausgeschlossen wurden. Ellefsen möchte aber in seiner weiteren Forschung diese genauer untersuchen, da sie zur Realität dazu gehören. Entwickelte Individuen benutzten entweder eine statische oder eine dynamische Plastizitäts-Regulierung und dafür waren verschiedene Anzahlen entwickelter Gene notwendig. Zwei der Gene wurden für die Codierung der angeborenen Strategien verwendet und die restlichen encodierten die Plastizität im Rest des Lebens des Individuums.

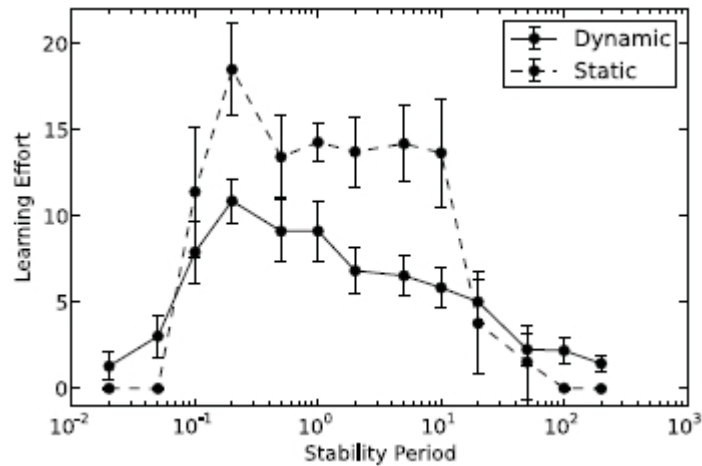


Abbildung 4: Die durchschnittlich erbrachte Lernleistung der Individuen über 20 Versuchsdurchführungen. Bei Dynamic kann die Lernleistung über die Lebensdauer variiert werden, bei Static ist diese immer gleich. Veränderungsrate sind angegeben in Generationen zwischen zwei Veränderungen. Die Fehlerbalken zeigen ein Konfidenzintervall von 95% um die Durchschnittswerte. Quelle: [6]

## 5 Ergebnisse

### 5.1 Veränderungsrate und Lernleistung

Um die Hypothese aus Abbildung 2 zu untersuchen, entwickelte Ellefsen die Lernleistung und angeborene Gewichtungen der Individuen unter verschiedenen Veränderungsrate der Umgebung. Diese Experimente wurden sowohl für Individuen durchgeführt, welche eine konstante Lernrate über die gesamte Lebensdauer haben, als auch für solche, deren Lernrate variiert.

Abbildung 4 zeigt die gemessene Lernleistung in verschiedenen Stabilitätsperioden. Diese ist proportional zu dem Lernaufwand, den ein Individuum in einem Zeitabschnitt gemacht hat. Bei einem dynamischen Lernaufwand entsprechend der dynamischen Plastizitäts-Regulierung waren vier Lernstrategien zu erkennen, welche der Hypothese aus Abbildung 2 entsprachen. Für Veränderungen, die zu langsam oder zu schnell auftreten, ist es nicht vorteilhaft, zu lernen. Dieses entspricht der Vermutung, dass das Lernen dem *Goldlöckchen-Prinzip* (siehe [11]) folgt. Die Änderungsrate muss sich also in einem perfekten Zustand befinden, damit als Reaktion auf Veränderung gelernt wird.

Wenn sich die Änderungsrate im benötigten Bereich befindet, damit die Individuen lernen, sind zwei Strategien bei den dynamischen Lernern zu beobachten. Dieser Bereich lag zwischen 0,1 bis 50 Lebenszeiten zwischen zwei Wechseln, also 0,02 bis 10 Veränderungen pro Generation. Die erste Strategie ist die, dass die Individuen ihr ganzes Leben plastisch sind. Diese Strategie ist in Umgebungen sinnvoll, in denen Veränderungen so schnell nacheinander auftreten, dass sensitive Phasen nicht ausreichen würden. Im Experiment

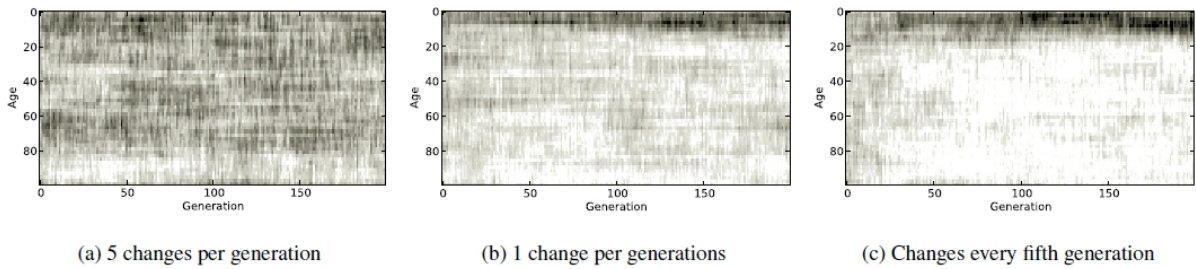


Abbildung 5: Die durchschnittliche Lernrate (aus 20 Versuchsdurchführungen) über die 100 Zeitschritte der Lebensdauer (Y-Achse) der besten Individuen aus jeder Generation der Evolution (X-Achse). Je dunkler, desto höher ist die Lernrate. Quelle: [6]

trat dieses Lernverhalten auf, wenn zwischen 1 und 10 Veränderungen pro Generation stattfanden. In Abbildung 5 erkennt man, dass bei einer Stabilitätsperiode von 1 ein Wendepunkt zu erkennen ist, an dem die untersuchten Individuen ihre Lernstrategie wechseln. Tritt weniger als eine Veränderung pro Generation auf, sind die Individuen nur zu Beginn ihres Lebens plastisch und danach nicht mehr lernfähig. Dadurch ergibt sich der Vorteil, dass sich an die aktuellen Gegebenheiten angepasst werden kann und die Kosten der lebenslangen Plastizität eingespart werden können.

Abbildung 6 zeigt die Lernleistung über die gesamte Lebensdauer und wie sie sich unter verschiedenen Veränderungsraten der Umgebung verändert. Die hellen Abschnitte am linken und rechten Rand weisen darauf hin, dass dort nicht gelernt wird, da hier die Veränderungen zu langsam oder zu schnell auftreten. Liegt die Stabilitätsperiode bei einem Wert zwischen 0,5 und 1, erkennt man an den durchgehend grauen Spalten, dass eine lebenslange Plastizität entwickelt wurde. Bei Werten zwischen 1 und 20 sind sensitive Phasen entwickelt worden. Diese Ergebnisse stützen die aufgestellte Hypothese aus Abbildung 2.

## 5.2 Statische oder dynamische Plastizität

Der Vergleich zwischen plastischen und statischen Individuen hilft dabei, die Wichtigkeit von sensitiven Phasen beim Abwägen von Lernkosten und -nutzen zu verdeutlichen.

Wie in Abbildung 4 zu erkennen ist, besteht der größte Unterschied zwischen den beiden Individuen-Typen darin, dass die dynamischen Lerner ihre Lernrate gleichmäßiger an unterschiedliche Umgebungsveränderungsraten anpassen können. Statische Lerner hingegen haben nur einen *on/off-Modus*. Sie können also entweder mit voller Leistung lernen oder gar nicht. Veränderungsrate, in denen sich für das Lernen entschieden wird, führen dazu, dass Individuen viel ins Lernen investieren. Aufgrund ihrer Unfähigkeit, die Lernleistung über ihre Lebensdauer anzupassen, müssen statische Lerner ihre Lernfähigkeit früher abschalten, damit die Lernkosten nicht den Lernnutzen übersteigen. Dynamische Lerner hingegen bleiben immer ein wenig lernfähig, auch wenn es unter

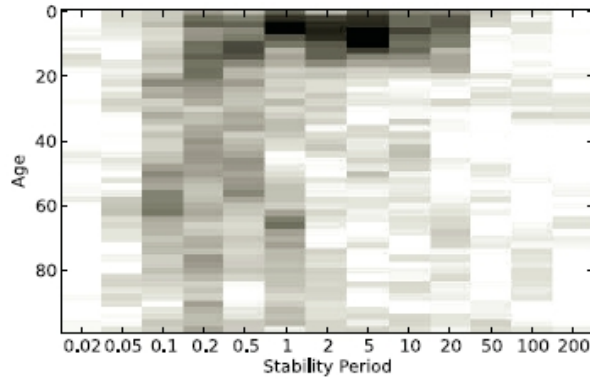


Abbildung 6: Der geleistete Lernaufwand der Individuen bei verschiedenen Umgebungsveränderungsraten. Dunkle Einträge weisen auf eine hohe Lernrate hin. Quelle: [6]

den gegebenen Bedingungen nicht nötig wäre. Sie fahren aber ihre Lernfähigkeit so weit herunter, dass die Kosten nicht zu hoch sind.

Eine weitere Möglichkeit, statische und dynamische Individuen zu vergleichen, bietet die Beobachtung der Fitnesswerte direkt vor und direkt nach einer Umgebungsveränderung. Dieser Vergleich zeigt, wie gut Individuen ihre Lernleistung ausbalancieren und dabei die Vorteile des Lernen ausnutzen können, aber Kosten dafür weitestgehend vermeiden. Abbildung 7(a) zeigt die Fitnesswerte der besten Individuen in einer Generation direkt vor einer Veränderung. Bei einer Stabilitätsperiode von 50 haben sich die letzten 49 Generationen unter gleichbleibenden Bedingungen entwickelt. Es ist zu sehen, dass dynamische Individuen einen deutlichen Fitnessanstieg vorweisen können, wenn eine Veränderung zwischen jeder zweiten und jeder zehnten Generationen auftritt. Dieses ist auch der Zeitraum, in der sensitive Phasen eine große Rolle spielen. Für schnellere Veränderungen ist lebenslange Plastizität vorteilhafter, wie es in Abbildung 5 gezeigt ist. Für langsamere Veränderungen ist lernen nicht mehr vorteilhaft, stattdessen passen sich die Individuen genetisch an und vermeiden somit die Lernkosten.

Anders verhält es sich in der Generation direkt nach einer Veränderung, wie in Abbildung 7(b) zu erkennen ist. Dynamische Lerner haben einen Vorteil gegenüber den statischen Lernern, wenn die Umgebung für lange Zeit stabil war. Das ist damit zu begründen, dass statische Individuen sich immer mehr auf die genetische Anpassung verlassen und nicht so schnell auf Veränderungen reagieren können. Die unterschiedliche Anpassungsfähigkeit innerhalb einer Art ist an den Fehlerbalken zu erkennen, welche teilweise sehr groß sind. Das ist eine natürliche Folge aus der Veränderung der Umgebung. Die vor der Veränderung am besten agierenden Individuen, versuchen ihre Handlungen weiterhin durchzuführen. Diese sind nun aber nach der Veränderung eventuell am ungeeignetsten und können so zu den größten Fehlern führen.

Der gleiche Effekt ist zu beobachten, wenn man die Fitnesskurven der Individuen ohne Lernkosten anschaut. Hierbei wird gemessen, wie groß der Anteil an verträglicher Nahrung an der Gesamtmenge der aufgenommenen Nahrung ist.

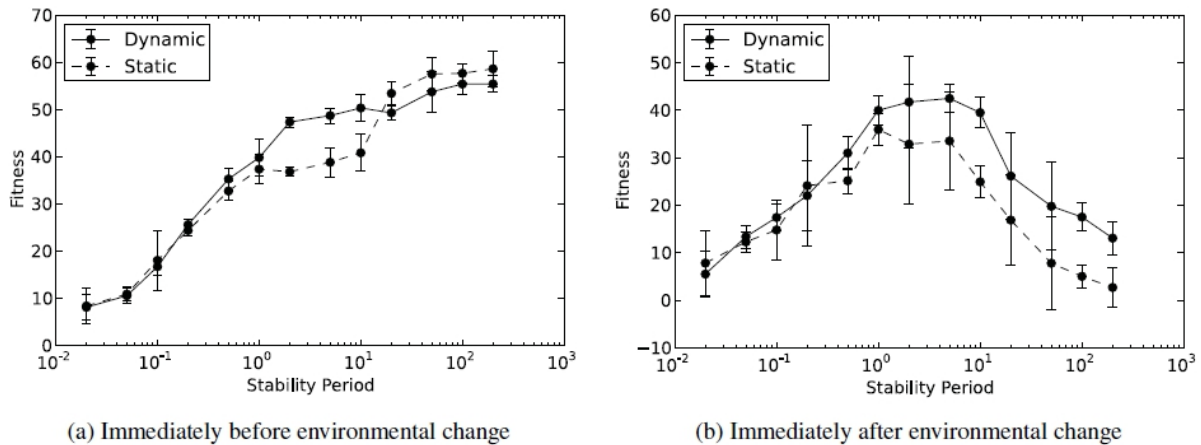


Abbildung 7: Die Fitnesswerte von den besten dynamischen und statischen Individuen im Vergleich. Gemessen wurde direkt vor und direkt nach einer Umgebungsveränderung. Die Fehlerbalken zeigen ein Konfidenzintervall von 95% um die Durchschnittswerte über 20 Versuchsdurchführungen. Quelle: [6]

Abbildung 8 zeigt die Entwicklung der Fitnesswerte über 200 Generationen. Die deutlichen Abfälle in den Kurven deuten auf die aufgetretenen Veränderungen in der Farbzuordnung der Nahrung hin. Hierbei fallen die statischen Individuen deutlich stärker ab als die dynamischen. Sie haben sich mehr auf die genetische Anpassung verlassen und reduzierten ihre Lernkapazitäten auf ein Minimum.

Auch dynamische Lerner haben ihren Lernaufwand über die Zeit deutlich verringert, dennoch liegt dieser über dem von den statischen Lernern. Sie können ihre Lernfähigkeit über Generationen erhalten, ohne dabei hohe Kosten zahlen zu müssen.

Bei mehr als einem Wechsel pro Generation sind die statischen Lerner jedoch oftmals effektiver als die dynamischen. Diese weisen dann eine lebenslange hohe Lernfähigkeit auf, wodurch sie den dynamischen gegenüber im Vorteil sind.

## 6 Schlussfolgerung

Bei der Untersuchung der Entwicklung von Lernstrategien bei verschiedenen Veränderungsrate der Umgebung wurden vier Hauptstrategien entdeckt. Bei zu niedrigen oder zu hohen Raten wurde kein Lernen von den untersuchten Individuen entwickelt, da die Kosten des Lernens höher sind als dessen Nutzen. Wenn die Veränderungsrate in dem Bereich liegt, in dem gelernt wird, werden zwei Strategien unterschieden. Zum einen gibt es die lebenslange Plastizität, welche in Umgebungen vorteilhaft ist, die sich schnell verändern. Zum anderen gibt es die Entwicklung von sensitiven Phasen für geringe Veränderungsrate.

Es gibt zwei wichtige Eigenschaften, die die dynamischen von den statischen Individuen unterscheiden. Erstens: Bei wenigen Veränderungen der Umgebung weisen die

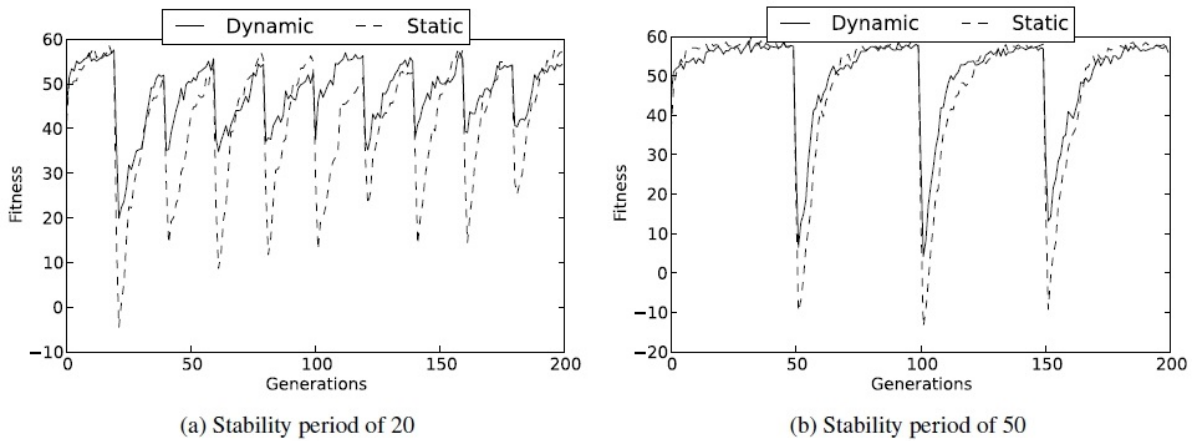


Abbildung 8: Fitnesswerte ohne Kosten für Plastizität. Quelle: [6]

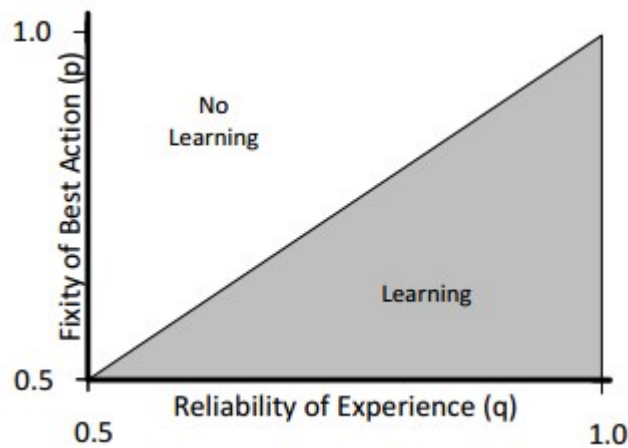


Abbildung 9: Das pq-Modell von Dunlap und Stephens. Quelle: [7]

dynamischen Individuen eine geringere genetische Anpassung auf und zweitens können diese denselben Nutzen durch lernen erzielen, während sie weniger Kosten dafür haben. Dafür müssen sie sensitive Phasen entwickeln.

Die Ergebnisse zeigen auf, dass genetische Anpassung und sensitive Phasen ähnliche Rollen einnehmen. Sie reduzieren die Lernkosten, während die Vorteile beibehalten werden. Dadurch, dass sie ähnliche Rollen spielen, nehmen sie aufeinander Einfluss. So reduzieren sensitive Phasen die Notwendigkeit für genetische Anpassung und umgekehrt. Deswegen ist es wichtig, dass beide Eigenschaften in einem Modell zusammen untersucht werden müssen, um die jeweiligen Rollen zu verstehen.

## 7 Weiterführende Forschung

Auf Grundlage seiner bisherigen Arbeiten erforschte Ellefsen weitere Aspekte der Entwicklung des Lernverhaltens unter verschiedenen Umwelteinflüssen. Im Jahr 2014 veröffentlichte er eine Arbeit, in der er die Agenten in drei verschiedenen Umgebungen aussetzte [7]. In der ersten Umgebung gab es eine lange Phase, in der Erfahrungen gesammelt werden konnten, auf die dann eine Veränderung der Umgebung folgte. In der zweiten Umgebung gab es im Wechsel eine Erfahrungsphase und eine Veränderungsphase. Die letzte Umgebung war immer im Wandel, zwischen Erfahrungsphasen und Veränderungsphasen konnte nicht unterschieden werden, da sie vermischt waren. Ellefsen griff auf das *pq-Modell* von Dunlap und Stephens zurück (siehe dazu [5]). Dieses Modell (Abbildung 9) stellte dar, bei welchen Werten der *best-action fixity* und *reliability of experience* sich für bzw. gegen lernen entschieden wird. Allerdings bezogen sie sich dabei nur auf das Umgebungsmodell, bei dem auf eine lange Erfahrungsphase eine Veränderung eintritt.

Ellefsen fand heraus, dass sich dieses Modell auch auf die realistischeren Umgebungen (zweites und drittes Modell) übertragen lässt.



## Literaturverzeichnis

- [1] Anderson, R. (1995). Learning and evolution: A quantitative genetics approach. *Journal of Theoretical Biology*, 175(1):89-101.
- [2] Baldwin, M. (1896). A New Factor in Evolution. *The American Naturalist*, (30):536-553.
- [3] Bullinaria, J. A. (2003). From biological models to the evolution of robot control systems. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 361(1811):2145-2164.
- [4] Downing, K. L. (2010). A Script-Based Approach to Evolving Neural Networks. In *Proceedings of the second Norwegian Artificial Intelligence Symposium*, pages 29-36.
- [5] Dunlap, A. S.; Stephens, D. W. (2009). Components of change in the evolution of learning and unlearned preference. *Proceedings. Biological sciences / The Royal Society*, 276(1670):3201-3208.
- [6] Ellefsen, K. O., (2013). Balancing the costs and benefits of learning ability. In *ECAL 2013: Proceedings of the 12th European Conference on the Synthesis and Simulation of Living Systems*, pages 292-299. MIT, Cambridge, USA.
- [7] Ellefsen, K. O., (2014). The Evolution of Learning Under Environmental Variability. In *ALIFE 14: Proceedings of the Fourteenth International Conference on the Synthesis and Simulation of Living Systems*, pages 649-656. MIT, Cambridge, USA.
- [8] Floreano, D.; Urzelai, J. (2001). Evolution of Plastic Control Networks. *Autonomous Robots*, 11(3):311-317.
- [9] Hinton, G.; Nowlan, S. (1987). How learning can guide evolution. *Complex systems*.
- [10] Hubel, D. H.; Wiesel, T. N. (1970). The period of susceptibility to the physiological effect of unilateral eye closure in kittens. *Journal of Physiology*, 206(2):419-436.
- [11] Kerr, B.; Feldman, M. (2003). Carving the Cognitive Niche: Optimal Learning Strategies in Homogenous and Heterogenous Environments. *Journal of Theoretical Biology*, 220(2):169-188.
- [12] Kirby, S.; Herford, J. R. (1997). The evolution of incremental learning: language, development and critical periods. Technical report, Language Evolution and Computation Research Unit, University of Edinburgh.
- [13] Knudsen, E. I. (2004). Sensitive Periods in the Development of the Brain and Behavior. *Journal of Cognitive Neuroscience*, 16(8):1412-1425.
- [14] Littman, M. (1995). Simulations Combining Evolution and Learning. In *Adaptive Individuals in Evolving Populations: Models and Algorithms: Santa Fe Institute Studies in the Sciences of Complexity*, pages 465-477. Addison-Wesley.

- [15] Mayley, G. (1996). Landscapes, Learning Costs and Genetic Assimilation. *Evolutionary Computation*, 4(3):213-234.
- [16] Mayley, G. (1996). The Evolutionary Cost of Learning. In *From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive behavior*, pages 458-467.
- [17] Nolfi, S.; Parisi, D.; Elman, J. L. (1994). Learning and Evolution in Neural Networks. *Adaptive Behavior*, 3(1):5-28.
- [18] Sasaki, T.; Tokoro, M. (1999). Evolving Learnable Neural Networks under Changing Environments with Various Rates of Inheritance of Acquired Characters: Comparison between Darwinian and Lamarckian Evolution. *Artificial Life*, 5(3):203-223.
- [19] Todd, P. M.; Miller, G. F. (1991). Exploring adaptive agency II: Simulating the evolution of associative learning. In *From Animals to Animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior*, pages 306-315, Cambridge, Ma. MIT Press/Bradford Books.
- [20] Turney, P. (1996). Myths and Legends of the Baldwin Effect. In *13th International Conference on Machine Learning, Workshop on Evolutionary Computation and Machine Learning*, pages 135-142, Bari, Italy.
- [21] Watson, J.R.; Geard, N.; Wiles, J. (2002). Stability and task complexity: a neural network model of genetic assimilation. In *Proceedings of the 8th International Conference on Artificial Life (Artificial Life VIII)*, pages 153-156. MIT Press.