



**HOCHSCHULE  
MITTWEIDA**

University of Applied Sciences

---

Fakultät Angewandte Computer- und Biowissenschaften

Professur Digitale Transformation und Angewandte Medieninformatik

## **Bachelorarbeit**

Analyse der Vergleichbarkeit von Balancing-Aspekten in  
verschiedenen Spielgenres

Tim Riemann

Mittweida, den 20. September 2023

**Erstprüfer:** Prof. Dr.-Ing. Christian Roschke

**Zweitprüfer:** M.Sc. Manuel Heizing

**Riemann, Tim**

Analyse der Vergleichbarkeit von Balancing-Aspekten in verschiedenen Spielgenres  
Bachelorarbeit, Fakultät Angewandte Computer- und Biowissenschaften  
Hochschule Mittweida — University of Applied Sciences, September 2023

## **Referat**

Die nachfolgende Arbeit beschäftigt sich mit dem Testen und Bewerten von zuvor gewählten Metriken über drei Spiele aus verschiedenen Genres. Das Ziel ist es zu erörtern, ob und bis zu welchem Grad Genre übergreifendes Balancing möglich ist.

**Name:** Riemann, Tim

**Studiengang:** Medieninformatik und Interaktives Entertainment

**Seminargruppe:** MI19w3-B

**English Title:** Analysing the comparability of balancing aspects in various game genres

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	Zielstellung . . . . .	1
1.2	Aufbau der folgenden Arbeit . . . . .	2
<b>2</b>	<b>Theoretische Grundlagen</b>	<b>3</b>
2.1	Was sind Spielgenre . . . . .	3
2.2	Genredefinitionen: Genreaspekte in Depth analysieren . . . . .	5
2.2.1	Rundenbasierte Strategiespiele . . . . .	5
2.2.2	Fighting Games . . . . .	6
2.2.3	Echtzeit Battle Simulatoren . . . . .	7
2.3	Balancing Grundlagen . . . . .	8
2.4	Betrachtung der bestehenden Anwendungen . . . . .	9
2.4.1	The Battle for Wesnoth . . . . .	9
2.4.2	Rivals of Aether . . . . .	10
2.4.3	Totally Accurate Battle Simulator . . . . .	12
<b>3</b>	<b>Analyse und Konzeptionierung</b>	<b>15</b>
3.1	Metrikdefinition . . . . .	15
3.2	Gewichtungsannahmen . . . . .	16
<b>4</b>	<b>Forschungsprozess</b>	<b>21</b>
4.1	Totally Accurate Battle Simulator Modifikation . . . . .	21
4.2	Battle for Wesnoth Modifikation . . . . .	24
4.3	Rivals of Aether Modifikation . . . . .	28
<b>5</b>	<b>Evaluation</b>	<b>31</b>

5.1	Ziel . . . . .	31
5.2	Untersuchungsdesign . . . . .	31
5.3	Durchführung . . . . .	33
5.4	Auswertung . . . . .	35
5.4.1	Totally Accurate Battle Simulator . . . . .	36
5.4.2	Rivals of Aether . . . . .	39
5.4.3	Battle for Wesnoth . . . . .	41
5.4.4	Vergleich der Genre . . . . .	43
<b>6</b>	<b>Schlussbetrachtung</b>	<b>49</b>
6.1	Zusammenfassung . . . . .	49
6.2	Ausblick . . . . .	52
	<b>Literaturverzeichnis</b>	<b>I</b>
	<b>A Analysedokumente</b>	<b>I</b>
A.1	Exceltabellen . . . . .	I

# 1. Einleitung

Diese Bachelorarbeit wird sich mit der Thematik des Balancings beschäftigen und untersuchen, inwieweit Game Designer bestimmte Metriken auch über ein Genre hinaus verwenden können.

Beim Balancing von Videospielen geht es darum, die Regeln und Parameter der Spiele zu bestimmen und dafür zu sorgen, dass es fair und wünschenswert zu spielen ist. Dabei sorgt Balancing dafür, dass alle Teile des Spiels so wie geplant und reibungslos miteinander funktionieren und das kein einzelnes Element die anderen obsolet macht.

Des Weiteren ist das Balancen von Spielen ein langer Prozess, welcher nicht nur am Anfang sondern während der gesamten Zeit vonstatten geht. [BAA<sup>+</sup>16, S.1] Es wird deutlich, dass es sich dabei also um einen großen und wichtigen Teil bei der Entwicklung von Spielen handelt, weshalb sich die Frage ergibt, ob man bestimmte Metriken und Werte festlegen kann und diese genreübergreifend auf Spiele anwenden und entsprechend verändern kann.

## 1.1. Zielstellung

Das Ziel dieser Bachelorarbeit ist es, zu ergründen, inwiefern Errungenschaften der Balancing Welt auf weitere Genre übertragen werden können. Dies würde vor allem kleinen Teams mit Vertragsarbeiten helfen, welche stets Spiele von unterschiedlichen Genres für diese Vertragsarbeiten entwickeln müssen. Ebenso soll diese Bachelorarbeit helfen, das Verständnis für Balancing zu vertiefen und die einzigartigen Aspekte der betrachteten Genre klar darstellen und zur tieferen Untersuchung zur Verfügung stellen. Dies wird durch das mit Hilfe von Non-Player-Charakteren beschleunigte Testen von verschiedenen Balancing Aspekten, welche genreübergreifend analysiert und bewertet werden, erreicht.

## 1.2. Aufbau der folgenden Arbeit

Die folgende Arbeit ist grundlegend in sechs Kapitel unterteilt. Im ersten Kapitel, der Einleitung, wird die Zielstellung dieser Arbeit erläutert, sowie deren Aufbau kurz beschrieben.

Darauf folgend befasst sich das zweite Kapitel mit den theoretischen Grundlagen. In diesem Kapitel werden die Grundlagen, auf welchen die Annahmen der Arbeit basieren, erläutert. Diese beinhalten Erklärungen zum Thema Spielgenre, Definitionen der, in der Bachelor Arbeit verwendeten, Spielgenre, Klassifikationen dieser und eine Erklärung der für das Balancing gewählten Aspekte.

Das dritte Kapitel befasst sich mit der Analyse und Konzeptionierung. In diesem Kapitel werden die genauen Metriken gewählt, welche anhand der modifizierten Balancingaspekte der Genre mithilfe der Siegeswahrscheinlichkeit der modifizierten CPU's betrachtet werden. Zudem wird auf das Anlegen der Annahmen, mit denen die Funde später evaluiert werden, eingegangen

Darauf folgt im vierten Kapitel die Erstellung, der Modifikationen der Spiele, anschaulich gemacht durch das verwenden von Bildern und Skriptbeispielen.

Das fünfte Kapitel befasst sich mit dem Testen der Annahmen. Jede vorbestimmte Metrik wird mit modifizierten Werten getestet und jeweils mit einem Test ohne Veränderungen, einem mit Bonusstärken und einem mit negativen Schwächungen überprüft. Die Siegesraten und genauen Modifikationen werden hierbei jeweils festgehalten und mit Diagrammen dargestellt. Zusätzlich, werden Funde des Kapitels evaluiert und genauer betrachtet. Diese Funde werden in diesem Kapitel mit den Annahmen des dritten Kapitels verglichen und erneut ausgewertet.

Das finale Kapitel beschäftigt sich mit der Zusammenfassung der Funde und einem Ausblick der darstellt, welchen Effekt diese Arbeit auf die Spieleindustrie haben kann.

## 2. Theoretische Grundlagen

Das folgende Kapitel befasst sich mit den theoretischen Grundlagen der Bachelorarbeit. Dies beinhaltet eine allgemeine Erklärung, was Spielgenre sind und weshalb diese etabliert wurden.

Darauf folgend werden die in dieser Arbeit betrachteten Genre definiert und die Aspekte dieser genauer betrachtet. Ebenso werden die Grundlagen des Balancings in diesem Kapitel beleuchtet und erläutert.

Am Ende des Kapitels werden die genauen Anwendungen, welche im Laufe der Arbeit modifiziert werden, genau betrachtet.

### 2.1. Was sind Spielgenre

Ein wichtiger Aspekt für Forscher und Endnutzer ist es, Produkte, wie zum Beispiel Videospiele, klassifizieren zu können. Eine möglichst genaue Klassifizierung kann dem Endnutzer ein Gefühl dafür geben, ob das Produkt den gewünschten Anforderungen entspricht. Wissenschaftler und Forscher können die Spiele mit Hilfe von konkreten Klassifizierungen genauer betrachten und mit ähnlichen Kategorien vergleichen. Auch für Marketingzwecke und die Presse ist es wichtig, die Produkte zu bewerten, um sie an die richtigen Endnutzer zu vermarkten. In der Spieleindustrie ist die meistgenutzte und allgemein bekannte Klassifizierung die des Genres. Genres können hierbei von sehr allgemein wie von *Action* über *Abenteuer* reichen oder genauer und klarer definiert sein, wie bei den Sub-Genres *Metroidvania* oder *Soulslike*.

Die Spieleindustrie und die Genres innerhalb dieser sind im permanenten Wandel, jedes neue Spiel verbindet die Aspekte von alten Spielen mit den neuen Ideen oder den Ideen anderer Genres und Spiele. Dadurch kommt es zu einer nicht linearen



Evolution [Ars09], welche bestehende Genres immer weiter verändert, bis es zur Schöpfung von neuen Sub-Genres kommt.

Dadurch wächst jedes Jahr die Varianz an Spielen und die Anzahl an Genres, welche im Alltag verwendet werden und mit steigender Beliebtheit werden diese zu allgemein anerkannten Genres. Die bereits zuvor erwähnten Genres *Metroidvania* und *Soulslike* sind erst innerhalb der letzten 20 Jahre aus speziellen Spielen entstanden. *Metroidvania* entstand aus den Spielereihen *Metroid* (1986-2021) und *Castlevania* (1986-2014) [RCL, S.519], während *Soulslike* aus der Spielereihe *Dark Souls* (2011-2018), so wie *Demon's Souls* (2009) und *Bloodborne* (2015) entstand.[Wei, S.23]

Mit Spielgenres lässt sich fast jeder Aspekt von Spielen beschreiben. Vom Grundbaustein wie bei *Action*, *Adventure* und *Simulation* als grobe Beschreibung, wie der Spielablauf vonstatten geht oder genaueren Beschreibungen, wie das Spiel gespielt wird, wie *First-Person Shooter*, *Platform Runner* oder *Tower-Defense*.

Ernest Adams definiert Spielgenre als:

„Genres are categories of games characterized by particular kinds of challenge, regardless of setting or game-world content.“ [Ada13, S.67]

Diese Aussage besagt, dass das Genre eines Spieles nur bestimmt, welche Herausforderungen auf den Spieler zukommen und nicht, wie diese mit Kunststrichtungen oder Welteninhalt *verpackt* sind. Spielgenres stehen hiermit im Gegensatz zu Filmgenres, in welchen das Genre laut Thomas Schatz ein Ausdruckspektrum für die Filmproduzenten und ein Erfahrungsspektrum für die Zuschauer definiert [Sch81, S.22].

Weiterhin beschreibt Herr Adams auch das Problem, dass es in der Kreativitätsexplosion der Mobilspiel- und Internetspielindustrie immer schwerer wird, Spiele eindeutig in ein Genre zu klassifizieren. So kommt es, dass Spiele nur noch selten in genau ein einziges Genre eingeordnet werden können. Ermöglicht wird dies durch die bereits erwähnten neuen Kreativitätseinflüsse und die Fortschritte der Technik, welche es erlauben, mehr Inhalt und damit Spielelemente und Herausforderungen auf den Endgeräten der Benutzer zu speichern und darzustellen. William Higinbotham, der Entwickler von *Tennis for Two*, konnte mit seiner damaligen Technik weniger darstellen als heutige Mobiltelefone und hatte weniger Rechenleistung und Speicherplatz als heutige Minicomputer wie zum Beispiel der Raspberry Pie.

## 2.2. Genredefinitionen: Genreaspekte in Depth analysieren

Das erste Spiel welches für diese Bachelorarbeit betrachtet und modifiziert wird ist *Battle for Wesnoth*, ein rundenbasiertes Free to Play Indie Strategiespiel. Beim zweiten betrachteten Spiel handelt es sich um *Rivals of Aether*, ein Indie Fighting Game, welches sich an die erfolgreiche Reihe Super Smash Brothers von Nintendo anlehnt. Abschließend wird *Totally Accurate Battle Simulator* betrachtet, ein Spiel welches ein neues Minigenre, das der Battle Simulatoren, bekannt machte und daher ein Echtzeit Battle Simulator, Indie, Simulator, Strategie Spiel ist.

Aus diesen drei Titeln bilden sich die betrachteten Hauptgenres: rundenbasiertes Strategiespiel, Fighting Game und Echtzeit Battle Simulator.

### 2.2.1. Rundenbasierte Strategiespiele

Rundenbasierte Strategiespiele sind ein Sub-Genre des großen Genres der Strategiespiele, welche vor allem dadurch klassifiziert werden, dass der Spieler herausgefordert wird durch Planung und genau geplante und gezielte Aktionsabläufe den Sieg gegen einen oder mehrere Gegner zu erreichen. [Ada13, S. 72]

Bei rundenbasierten Strategiespielen bleibt dem Spieler durch das rundenbasierte Design mehr Zeit und dadurch bessere Chancen Entscheidungen zu treffen und die potentiellen Bewegungen des Gegners und der eigenen Einheiten abzuwägen. Dies ähnelt sehr dem System von Schach; in beiden Systemen führt taktisches Denken zum Sieg. Der Unterschied zu Schach ist jedoch, dass nicht nur die Positionierung, sondern auch ein wenig Glück und ein gewisses Schere-Stein-Papier-System zum Zuge kommt.

Rundenbasierte Strategiespiele arbeiten oft mit einem Zufallssystem, welches dafür sorgt, dass keine Einheit absolute Kontrolle über das Feld erlangen kann. Battle for Wesnoth zum Beispiel arbeitet mit prozentualen Trefferwahrscheinlichkeiten, welche in der Regel nicht vom Angreifer, sondern von der Verteidigungswertung des Verteidigers auf dessen aktuellem Terrain bestimmt werden. So hat eine Ratte die gleiche Chance, einen Zwerg auf einem Berg zu treffen wie ein Drache, welcher versucht, einen Zwerg auf einem Berg zu treffen. Diese Aspekte bieten dem zurückliegenden

Spieler eine Chance, die Oberhand zu gewinnen. Zusätzlich spielen sich rundenbasierte Strategiespiele meist auf Karten mit Ressourcenpunkten und Ausbildungsstätten ab. Die Einheiten der Spielenden können diese erobern und erhalten jede Runde Ressourcen oder können Einheiten rekrutieren.

Battle for Wesnoth war das Grundspiel, von dem aus die Genres und anderen Spiele definiert wurden. Die Entscheidung fiel auf Battle for Wesnoth, da dieses Spiel einerseits sehr modifizierbar ist und andererseits, weil das rundenbasierte Strategiespiel Genre eine Vielzahl an unterschiedlichen Unteraspekten bietet, welche für den weiteren Verlauf der Arbeit modifiziert werden können. Mit dem Genre der rundenbasierten Strategiespiele als Forschungsgrundlage wurden Grundparameter von Lebenspunkten, Schaden und potentiell von Angriffsgeschwindigkeit definiert.

### 2.2.2. Fighting Games

Das Fighting Game Genre ist ein Genre mit großem Fokus auf Einzelkämpfe von einem Spieler gegen einen anderen Spieler oder den Computer. Zur Verfügung stehen beiden Kämpfern hierbei eine Vielzahl an spielbaren Charakteren, welche je den Spielstil und die visuellen Aspekte der Kämpfe verändern und teils komplett andere Fähigkeiten und Combos bieten. Das Ziel eines Fighting Games ist es, die Lebenspunkte des Gegners auf null zu bringen oder in der speziellen Untervariante der *Smash Games*, zu welchen *Rivals of Aether* gehört, den Gegner von der Plattform zu schlagen. Ein Beispiel im Sport hierfür ist Sumoringen, an welches *Smash Games* leicht angelehnt sind.

Fighting Games besitzen zusätzlich noch eine Menge an Kombinationsmöglichkeiten und haben von Haus aus eine sehr hohe Messlatte an Spieler-Skill-Relevanz. Dies führt zur großen, mit Fighting Games verbundenen, kompetitiven Szene. [Ket] Fighting Games leben das Motto von *Easy to learn, hard to master* voll und ganz.

Das Fighting Game Genre wurde gewählt, um einen großen gameplaymäßigen und genremäßigen Unterschied zu den anderen Spielen der Bachelorarbeit zu bieten. Fighting Games sind vor allem durch ihre Echtzeit-Kämpfe mit großer Möglichkeit der Verbesserung der Spielerfähigkeiten und einem direkten Einfluss dieser bekannt. Dieser große Einfluss im Gameplay bietet somit auch einen Unterschied in der Balancingrelevanz. Durch die Vergrößerung der Effekte von Spielerfähigkeiten hebt sich erneut der Vorteil der durch künstliche Intelligenz durchgeführten Tests hervor.

### 2.2.3. Echtzeit Battle Simulatoren

Echtzeit Battle Simulatoren sind ein kleines Subgenre der Echtzeit Strategiespiele. Bei diesem Subgenre hat der Spieler während des Kampfes keinen weiteren Einfluss auf dieses. Spiele dieses Genres sind bekannt für die Möglichkeit, immense Schlachten mit hunderten von kämpfenden Einheiten zu simulieren. *Totally Accurate Battle Simulator* war eines der ersten dieser Spiele, welches bereits 2016 die *ulkige Freude von Spielzeugfiguren zum Leben erweckt*. [Gar16]

Dies wird durch das Platzieren von verschiedensten Einheiten auf einem Schlachtfeld erreicht. Diese Einheiten kämpfen dann ohne den Einfluss des Spielers gegen ihre Gegner. Der Spieler wird hierbei innerhalb von Missionen durch vorgegebene Geldressourcen begrenzt und je nach Mission zusätzlich durch einen begrenzten Einheitenpool. Außerhalb von Missionen wird der Spieler lediglich durch seine Kreativität oder die Stärke seines Computers zurückgehalten. Das Ziel von Echtzeit Battle Simulatoren ist es, die gegnerische Armee zu besiegen.

Der primäre Charme dieses Genres ist die Mischung aus Taktik und absolutem Kontrollverlust, sobald der Kampf startet. Einerseits wird dem Spieler so die Macht gegeben, seine Armee zusammenzustellen und das perfekte Team zu finden. Sobald jedoch der Kampf beginnt, kann der Spieler sich komplett auf den Kampf konzentrieren und die einzelnen Kämpfe genießen, ohne den Hintergedanken zu haben, dass seine Flanke überrannt wird sobald diese nicht gemicromanaged wird. Dies wäre der Fall in anderen Echtzeit Strategiespielen wie Spielen der Reihe *Total War* oder *Wargame: Red Dragon*. Battle Simulatoren bieten somit epische cinematische Kamerafahrten erster Hand ebenso wie Katharsis durch teils ulkige und unfassbar unrealistische Kämpfe.

Das Genre der Echtzeit Battle Simulatoren wurde für diese Bachelorarbeit und deren Tests gewählt, um eine weitere Konstante zum Testen hinzuzufügen. Die Echtzeit Battle Simulatoren verfügen hierbei über Ähnlichkeiten zu beiden Genre. Einerseits ist das Kampfsystem ähnlich dem des rundenbasierten Strategie Spiels und basiert primär auf strategischem Können, aber gleichzeitig bietet der Echtzeit Battle Simulator einen Kampf, der ebenso wie das Fighting Game, in Echtzeit abläuft und von der vorhandenen Physik beeinflusst wird. Der primäre Unterschied zu den anderen beiden Genres ist hierbei jedoch, dass der Spieler keinen Einfluss auf den Kampf hat sobald dieser die Einheiten gewählt und den Kampf gestartet hat.

### 2.3. Balancing Grundlagen

Das Balancing der Spielmechaniken wird durch über Tests gestützte Annahmen und Anpassungen vonstatten gehen. Das grundlegende Ziel von Balancing in Spielen ist es, die Chancen von allen Parteien anzupassen.

Durch erhöhte Komplexität in Spielen steigt auch der Anspruch an das Balancing-department, da trotz erhöhter Komplexität und einer Vielzahl an unterschiedlichen Interaktionen zwischen Fähigkeiten und Aktionen das Ziel darin besteht, die Chancen aller Spieler anzugleichen, solange diese ähnliche Grundlagen an Fähigkeiten mitbringen.

Eine Analyse durch Becker und Görlich, durchgeführt im Jahr 2020, von 14 unterschiedlichen Balancing Konzepten ergab, dass diese sich zum größten Teil auf vier Aspekte fokussierten: den Charakteristiken von gut balancierten Spielen, der Schwierigkeit, der Symmetrie und dem Balancierungsprozess. [BG20]

Das Ziel des Balancing ist es, anhand dieser Aspekte den Spielspaß des Spieles zu maximieren. Im Laufe dieser Arbeit liegt der Fokus vor allem auf den Aspekten der Symmetrie und der Schwierigkeit über die Genres hinweg.

Das Ziel ist es zu betrachten, ob die Anpassungen in einem Spiel einen ähnlichen Einfluss mit gleichen Anpassungen in anderen Genres erweisen. Grundlegend werden bestimmte Metriken vordefiniert und diese für jedes Genre gleichmäßig angepasst.

Die Effektbewertung der Anpassungen der Metriken geschieht dann durch das Betrachten der veränderten Siegeswahrscheinlichkeiten. Eine Betrachtung der Siegeswahrscheinlichkeit von Kämpfern, welche immer auf dem gleichen Fähigkeitsgrad bleiben, ermöglicht es hierbei genau zu analysieren welchen Einfluss die modifizierten Metriken haben.

Durch die Verwendung der in den Spielen gegebenen künstlichen Intelligenz ist es hier möglich zu garantieren, dass sich die Fähigkeitsgrundlagen der Parteien nicht während der Tests verändern. Zusätzlich ermöglicht die Verwendung der künstlichen Intelligenz, mehrere Tests gleichzeitig mit identischen Bedingungen und Testern ablaufen zu lassen, ohne potentielle Ermüdung von Testern oder mehrere unterschiedliche Tester und deren Zeit zu benötigen.

Zum Testen dieser Balancing Grundlagen mithilfe der künstlichen Intelligenz wurden zuvor Recherchen betrieben um festzustellen, wie aussagekräftig durch künstliche Intelligenz durchgeführte Balancing Tests sind. Alexander Benjaming Jaffe schrieb hierzu 2013:

„This is where automated analysis may have a tremendous impact: in finding those imbalances that we do not need human beings to discover, leaving humans as much time as possible to express their uniqueness in the games that we do playtest.“ [Jaf, S. 65f.]

Dies bedeutet für diese Arbeit, dass die Tests der künstlichen Intelligenz für diese Art von Balancing Testing mehr als ausreichend sind, da mit dieser nur rein quantitative Daten erhoben werden, für welche die Kreativität der Spieler tatsächlich negativ sein könnte.

## 2.4. Betrachtung der bestehenden Anwendungen

In diesem Kapitel werden die Anwendungen, welche für diese Arbeit modifiziert werden und als Forschungsgrundlage dienen, genauer beschrieben und betrachtet.

### 2.4.1. The Battle for Wesnoth

The Battle for Wesnoth ist laut der offiziellen Website ein Open Source rundenbasiertes Strategie Spiel mit einem *High Fantasy* Thema, welches auch als *Epische Fantasy* bekannt ist. The Battle for Wesnoth erschien am 18. Juni 2003 und erhält bis heute monatlich neue Updates.

Open Source Software ist dafür bekannt, dass deren Quellcode offen verfügbar ist und ohne die Erlaubnis des Besitzers eingesehen und frei verwendbar, modifizierbar und teilweise sogar verteilbar ist.

Aus diesem Grund ist Open Source Software perfekt für diese Arbeit geeignet, da die Anwendungen für die geplanten Untersuchungen modifiziert werden müssen und mithilfe der vorhandenen Foren Fehler und Ansätze schneller gefunden werden können um die Modifikationen umzusetzen.

Der Fokus der Entwickler steht auf Singleplayer und Co-op-Kampagnen und somit dem PVE, also *Player versus Environment*. Dies bedeutet, dass der Spieler, mit anderen oder alleine, gegen die Gefahren des Spiels kämpft ohne einen Spieler als Gegner zu haben. Auch der PVP- Modus, also *Player versus Player* beziehungsweise der Kampf von Spieler gegen Spieler, wird vom Spiel auf den offiziellen Servern unterstützt.

Aufgrund des Open Source Aspekts gibt es auch kleinere Modifikationen, in denen eine kompetitive Szene namens *Afterlive Balanced* von der Spieler Community erstellt wurde und PvP-Server zur Verfügung steht.



Abbildung 2.1.: Ein Kampf auf einer zufällig generierten Karte in Battle for Wesnoth

### 2.4.2. Rivals of Aether

Rivals of Aether ist laut der offiziellen Website ein *Indie Fighting* Spiel, welches in einer Welt spielt, in welcher Zivilisationen mithilfe von dem Heraufbeschwören von Feuer, Wasser, Luft und Erde Krieg führen. Die Spieler kämpfen mit sogenannten Rivalen, welche sie in den Kampf schicken und die Elemente kontrollieren lassen. Besonders beliebt ist es für das kompetitive Gameplay und die intensiven Trainingsmöglichkeiten, was es zu einem Einstiegsspiel in das Fighter Plattform Genre macht. [Stuc]

Rivals of Aether wurde von Aether Studios veröffentlicht und gepublished. Die Entwicklung begann im April 2014 und wurde im Juli desselben Jahres öffentlich angekündigt. Nach drei Jahren Entwicklung wurde es schließlich am 28. März 2017 für PC und Xbox One veröffentlicht. [Stua]

Rivals of Aether bietet seinen Spielern verschiedene Modi. Im Versus Modus können Spieler einzeln, in Zweiertteams oder in einem Alle-Gegen-Alle Modus gegeneinander antreten. Diese Online Kämpfe können auch im sogenannten Ranked Modus ausgetragen werden, bei denen die Spieler um die Ranglistenplätze kämpfen können.

Der Abyss Modus stellt einen wellenbasierten Überlebenskampf dar, in dem die Spieler einzeln oder zu zweit gegen Wellen an Gegnern möglichst lang überleben müssen. Dort besteht die Möglichkeit, den Kampf individuell anzupassen und Freunde herauszufordern.

Auch für Spieler, die es etwas ruhiger angehen wollen, ist gesorgt. Im Story Modus wird die Hintergrundgeschichte der einzelnen Rivalen sowie die Geheimnisse des Aethers aufgedeckt. Dieser bietet auch die Möglichkeit zum lokalen Mehrspielermodus. Für zusätzliche Abwechslung sorgt der Tetherball Modus, ein sportliches Spiel im Strandvolleyball Stil.

Für Spieler, die neu lernen oder ihre Fähigkeiten verbessern wollen bietet Rivals of Aether einen Tutorial Modus, welcher ihnen die Möglichkeit gibt zu üben und Mechaniken zu verfeinern. Im Übungsmodus werden zudem noch nützliche Daten und Hitboxen der Spieler und Gegner visualisiert, um die eigenen Kämpfe und Fähigkeiten bestens zu analysieren. [Stub]



Abbildung 2.2.: Ein Kampf im Spiel Rivals of Aether



### 2.4.3. Totally Accurate Battle Simulator

Totally Accurate Battle Simulator wird auf der offiziellen Website beschrieben als eine Möglichkeit, den Leiter einer Armee von roten oder blauen *Wobblern* zu verkörpern. Diese sind wackelige und alberne Menschen von antiken Ländern, gruseligen Orten und fantastischen Welten. Der Spieler wählt in diesem Spiel seine Armee von Einheiten und betrachtet deren Kämpfe in Simulationen mit dem laut der Website wackeligsten Physiksystem der Welt.[Gam]

Das Spiel setzt hierbei einen Fokus auf den Sandboxmodus, in welchem man die Einheiten beider Seiten komplett frei wählen kann. Im Laufe der Entwicklung kamen zum Sandboxmodus noch viele weitere Funktionen hinzu; eine Kampagne mit zwölf Welten, die Möglichkeit eigene Karten und Einheiten zu erstellen, eine eigene Workshopeinbindung, mit welcher Spieler ihre eigenen Erstellungen mit anderen Spielern teilen können und Online Kämpfe gegen andere Spieler weltweit.



Abbildung 2.3.: Eine Nahaufnahme einer Schlacht in Totally Accurate Battle Simulator

Totally Accurate Battle Simulator begann seine Erfolgsgeschichte im Jahr 2016 mit 14 Einheiten und einer einzigen spielbaren Karte. Heute besitzt Totally Accurate Battle Simulator, ohne Modifikationen aus dem ingame Workshop, 139 Einheiten und 44 Karten. Mit der Hilfe vom Einheiten- und Kartenerstellungssystem wurden bereits mehr als 100.000 Szenarios, Einheiten, Fraktionen und Karten erstellt.

Ein wichtiges Ausschlusskriterium, das alle Spiele dieser Arbeit erfüllen mussten, war das der Modifikation. Auch *Totally Accurate Battle Simulator* erfüllt dieses Kriterium durch dessen Einheitenerstellungsfunktion. Diese Funktion erlaubt es, das Spiel und die darin vorhandenen Einheiten spielend leicht und mit geringem Zeitaufwand nach den Bedürfnissen der Tests zu modifizieren.



## 3. Analyse und Konzeptionierung

Dieses Kapitel dient der theoretischen Vorbereitung der, für die Evaluationsdurchführung benötigten, Metriken und der Modifikationen der Anwendungen.

Das Definieren der in den Tests betrachteten Metriken für die in der Bachelorarbeit vertretenen Genre erfolgt im ersten Unterkapitel. Im zweiten Unterkapitel werden Annahmen dazu getroffen, welchen Einfluss die Modifikationen der Metriken auf die prozentuale Siegesrate der modifizierten Einheiten haben werden.

### 3.1. Metrikendefinition

Die Grundmetriken, die innerhalb dieser Arbeit betrachtet werden, sind Schaden, Angriffsreichweite, Lebenspunkte und Angriffsgeschwindigkeit.

Die erste betrachtete Metrik dieser Arbeit ist die der Schadenspunkte. Schadenspunkte sind in Spielen Werte, die bestimmen wie viele Lebenspunkte eine Einheit verliert nachdem die aktive Einheit einen erfolgreichen Angriff gegen diese ausgeführt hat.

Angriffsreichweite ist die Metrik, welche bestimmt auf welche Distanz eine Einheit Angriffe gegen andere ausführen kann. Eine hohe Angriffsreichweite erlaubt es der Einheit andere zu treffen ohne selbst in Gefahr zu kommen.

Die dritte Metrik ist die der Lebenspunkte. Die Lebenspunkte einer Einheit bestimmen, wie viel Schaden diese nehmen kann bevor sie aus dem Kampf austritt.

Die finale für diese Bachelorarbeit betrachtete Metrik ist die der Angriffsgeschwindigkeit. Die Angriffsgeschwindigkeit einer Einheit beschreibt wie häufig diese Angriffe durchführen kann. Eine hohe Angriffsgeschwindigkeit resultiert in einer großen Menge an Angriffen.

Die festgelegten Werte werden für diese Arbeit je vier mal getestet, je einen Test für jede Modifikation. Die Modifikation der Testwerte sind wie folgend; 70 Prozent, 100 Prozent, 110 Prozent, 130 Prozent und 150 Prozent.

Für die so eben genannten Prozente existiert eine Ausnahme, die Metrik der Reichweite. Diese Ausnahme kommt durch die Begrenzungen, welche Totally Accurate Battle Simulator mit sich bringt. Es gibt in diesem Spiel nur wenige Waffen, welche auf unterschiedlichen Reichweiten gleiche oder zumindest vergleichbare Angriffsmuster besitzen. Für die Tests wurden fünf Waffen mit ähnlichen Angriffsmustern und unterschiedlichen Reichweiten gefunden. Die Prozentwerte der Reichweite sind daher 75 Prozent, 125 Prozent, 150 Prozent und 200 Prozent.

Die anderen drei Metriken Schaden, Angriffsgeschwindigkeit und Lebenspunkte werden mit den zuvor genannten 70 Prozent, 110 Prozent, 130 Prozent und 150 Prozent modifiziert.

## 3.2. Gewichtungsannahmen

Im folgenden Unterkapitel werden Annahmen getroffen, wie sich die Modifikationen der vorher festgelegten Metriken auf das Spiel und somit auf die Siegesrate der modifizierten Einheiten auswirken.

### Lebenspunkte

Die Gewichtungsannahmen zur ersten Metrik, der Metrik der Lebenspunkte, sind wie folgend:

Es wird davon ausgegangen, dass bei 100 Prozent der Leben 50 von 100 Kämpfen gewonnen werden. Diese Annahme kommt daher, da beide Einheiten die selben Werte haben. Sie werden in der Wahrscheinlichkeitsberechnung daher als gleichwertig betrachtet, da deren Stärken identisch sind. Diese Tests werden pro Spiel nur einmal mit 100 Testdurchläufen durchgeführt, da 100 Prozent der Lebenspunkte identisch dem Test der 100 Prozent der anderen Metriken sind.

Im Testfall mit 70 Prozent der Leben jedoch wird davon ausgegangen, dass die Chancen zu gewinnen geringer sind, als wenn die vorherige Rechnung mit 70 Prozent der Stärke, also  $\frac{3}{7}$  betrachtet wird. Drei Siebtel sind umgerechnet  $0,4286(\dots)$ . Es wird also davon ausgegangen, dass die modifizierte Einheit weniger als 43 Kämpfe

gewinnen wird. Der Wert der  $3/7$  kommt daher, dass der modifizierte Wert 70 Prozent der Stärke des unmodifizierten beträgt.

Während bei 150 Prozent mehr als 60 Prozent der Kämpfe, also mehr als  $3/5$ , gewonnen werden und sich diese Zahl eher 70 Prozent annähert. Diese Annahme erfolgt aus dem Schluss, dass eine Einheit mit mehr Lebenspunkten im Umkehrschluss mehr Schläge gegen den Gegner ausführen kann, bevor diese selbst besiegt ist.

Bei 110 und 130 Prozent wird je erwartet, dass diese mit einer Rate von 55 Prozent und rund 65 Prozent gewinnen, dies ergibt sich wieder aus einer Erhöhung der errechneten Werte 52,3810 und 56,5217.

Diese Annahmen hängen jedoch bei Lebenspunkten und dem verursachten Schaden von den in Spielen bekannten Breakpoints ab. Hiermit sind die Punkte gemeint, die einen Unterschied bei festen Zahlen machen. Wenn zwei Einheiten mit je 100 Leben und 25 Schaden gegeneinander kämpfen, so wird die Einheit mit 101 Leben einen weiteren Schlag genau so gut aushalten, wie eine Einheit mit 125 Leben. Sobald der Breakpoint erreicht ist, sind jegliche Werte bis zum nächsten in einer Testumgebung ohne Varianz irrelevant. Die Umgebung mit der geringsten Varianz ist die von Battle for Wesnoth, da diese als einzige Varianz die Trefferwahrscheinlichkeit bietet.

Rivals of Aether besitzt die Eigenheit, dass es in diesem Spiel keine Lebenspunkte gibt. Diese Metrik wurde daher bei Rivals of Aether ausgelassen.

### **Schaden**

Die Annahmen der Breakpoints zählen wie bereits erwähnt auch für den Schaden. Begeben wir uns zurück zu dem Beispiel mit 100 Leben: wird der Schaden einer der Einheiten um eins verringert und besitzt 24 Schaden, so braucht diese einen weiteren Schlag. Der Einheit ist es hierbei aber unwichtig, ob diese ein oder fünf Schaden verloren hat, da 20 den Breakpoint für fünf Schläge darstellt und 25 den für vier. Jeglicher Wert zwischen 20 und 25 ist irrelevant, das gleiche zählt auch nach oben. Erhält die Einheit einen Bonus von acht Schaden braucht sie weiterhin vier Schläge. Nach drei Schlägen ist der Gegner zwar auf einem Leben, jedoch ist dies in der varianzlosen Testumgebung gleichwertig mit 25 Leben. Erst der neunte weitere Bonusschaden verringert die benötigten Schläge erheblich von vier auf drei.

Die Testumgebungen der Spiele selbst bieten Varianz durch die Trefferwahrscheinlichkeit in Battle for Wesnoth, die ulkige Physik und Schadensvarianz in Totally Accurate Battlesimulator und das CPU Movement, random attack pattern und Dodge im Falle von Rivals of Aether.

Diese Eigenheit der Künstlichen Intelligenz in Rivals of Aether, und die Möglichkeit dieser, Angriffe mit geringer Rückstosskraft zu wählen obwohl der Gegner bereits bereit wäre von der Karte gestoßen zu werden, könnte den Einfluss von Schaden verringern. Um dies zu umgehen werden die Tests auf der höchsten CPU-Stärke durchgeführt um sicher zu gehen, dass diese die bestmöglichen Aktionen durchführen.

Deshalb wird davon ausgegangen, dass die Erfolgchancen für Battle for Wesnoth bei 10 Prozent Bonusschaden exakt 50 / 50 sind und für die anderen Werte identisch denen der HP sind. Für Totally Accurate Battle Simulator und Rivals of Aether wird davon ausgegangen, dass alle Werte denen der HP ähneln. Also weniger als 40 Prozent für die 70 Prozent Modifikation und 55, 60 und 70 Prozent für die 110, 130 und 150 Prozent Modifikationen.

#### **Angriffsgeschwindigkeit**

Im Falle der Angriffsgeschwindigkeit wird angenommen, dass diese sich ähnlich verhält wie der Schaden und die Lebenspunkte, jedoch mit unterschiedlichen Effekten in den Genres. Es wird davon ausgegangen, dass im Spiel Totally Accurate Battle Simulator der Effekt von Angriffsgeschwindigkeit erhöht ist, da eine Einheit mit hoher Angriffsgeschwindigkeit den Gegner mithilfe der ulkigen Umgebung davon abhalten kann anzugreifen. Bei den Spielen Rivals of Aether und Battle for Wesnoth wird davon ausgegangen, dass der Effekt der Angriffsgeschwindigkeit ähnlich bis identisch zu dem der Lebenspunkte ist.

Rivals of Aether birgt hierbei jedoch wieder eine Eigenheit. Es kann dazu kommen, dass die Angriffsgeschwindigkeit die künstliche Intelligenz verwirrt. Dabei kann es vorkommen, dass diese bei einer extrem langsamem Angriffsgeschwindigkeit zu langsam ist um zu reagieren, aber gleichzeitig eine so lange Angriffsdauer hat, dass die andere Einheit in den Schaden hineinläuft. Dies könnte potentiell die Werte ausgleichen und extreme Ergebnisse verhindern.

Es wird also angenommen, dass die Werte von Battle for Wesnoth ähnlich denen der Lebenspunkte sind, die Werte von Rivals of Aether stärker normiert, also sich stärker an ein 50 zu 50 Verhältnis annähern und die Werte von Totally Accurate Battle Simulator extremer sind.

Daraus ergeben sich die folgenden Annahmen: es wird davon ausgegangen, dass jeweils weniger als 40, 55, 60 und 70 Prozent gewonnen werden für die Modifikationen der 70, 110, 130 und 150 Prozent Werte. Bei Rivals of Aether ergeben sich rund 43, rund 53, rund 58 und rund 65 Prozent und bei Totally Accurate Battle Simulator weniger als 35, 60, 65 und 80 Prozent.

### **Angriffsreichweite**

Die Angriffsreichweite ist für die Vorbewertung ungewiss, da der Effekt von Spiel zu Spiel extrem volatil sein könnte. Es wird davon ausgegangen, dass der Effekt der Angriffsreichweite einen stärkeren Einfluss auf die Siegeswahrscheinlichkeit haben wird als alle anderen Werte. Vor allem in Battle for Wesnoth, in welchem die Reichweite durch die modifizierte Trefferwahrscheinlichkeit ersetzt wird. Die Trefferwahrscheinlichkeit besitzt ab bestimmten, entscheidenden Punkten die Eigenheit immer und immer effektiver zu werden. Die Trefferwahrscheinlichkeit von 20 Prozent auf zehn Prozent zu verringern halbiert die Trefferwahrscheinlichkeit, während die Trefferwahrscheinlichkeit bei 100 Prozent Genauigkeit um die gleichen zehn Prozent zu verringern den Schaden nur um ein Zehntel verringert.

Es wird also davon ausgegangen, dass die Effekte bei der Reichweite noch stärker erkennbar sind als bei den Lebenspunkten. Wenn man zu der Rechnung der Lebenspunkte zurück geht wird davon ausgegangen, dass sich der Effekt im Testfall der 75 Prozent der Reichweite so auswirkt, dass die Siegesrate der modifizierten Einheit weniger als 40 Kämpfe - wahrscheinlich sogar weniger als 35 Kämpfe gewinnt.

Im Gegensatz dazu wird davon ausgegangen, dass im Falle der 200 Prozent mehr als 70 Prozent aller Kämpfe gewonnen werden und sich die Siegeswahrscheinlichkeit der modifizierten Einheit in der Umgebung der 80 Prozent befindet. Für die Fälle 125 und 150 Prozent Modifikationen werden Siegeschancen in der 60 und 65 Prozent Umgebung prognostiziert.





## 4. Forschungsprozess

In diesem Kapitel werden die Prozesse dokumentiert mit welchem die Modifikationen der Spiele programmiert und erstellt werden. Dies wird mit Codebeispielen und Applikationsscreenshots festgehalten.

### 4.1. Totally Accurate Battle Simulator Modifikation

Für die Modifikation der Anwendung *Totally Accurate Battle Simulator* war es nicht nötig in den Code der Anwendung einzugreifen, da es in der Anwendung selbst einen Einheitenersteller gibt. (Vergl. Abbildung 4.1)

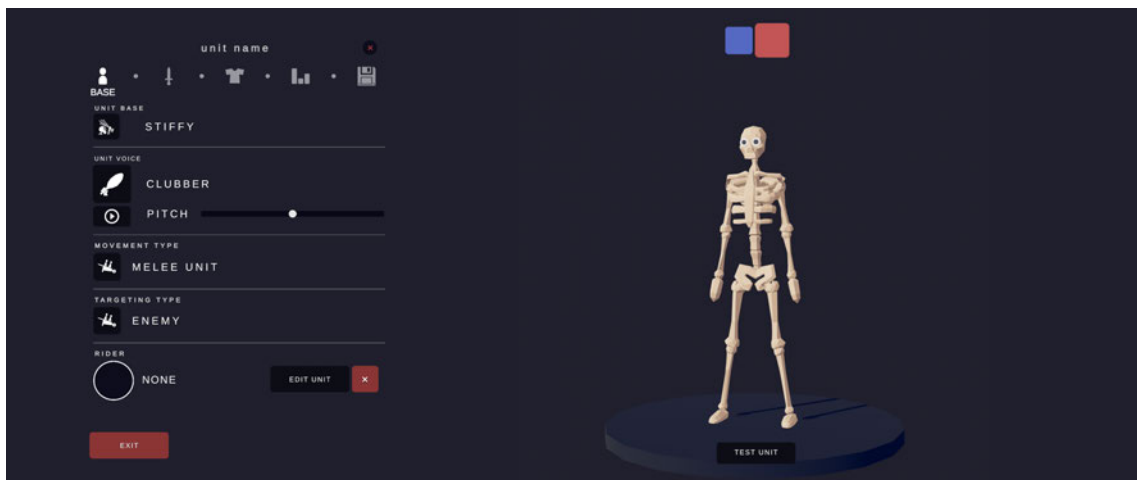


Abbildung 4.1.: Der Einheitenersteller von Totally Accurate Battle Simulator

Der Einheitenersteller von Totally Accurate Battle Simulator besitzt eine sehr große Varianz bezüglich aller Aspekte des Erstellens.

Auf dem ersten Bildschirm kann der Spieler die Wahl der Einheitenbasis treffen. Hierbei gibt es drei Wahlmöglichkeiten; Wobbler, Halfling und Stiffy. Darunter kann der Spieler sich entscheiden welche Stimme die Einheit besitzen soll und wie diese von der Tonhöhe her angepasst werden soll. Danach bestimmt der Spieler das Bewegungsmuster, das Zielverhalten und ob die Einheit eine andere Einheit auf ihren Schultern tragen soll.

Der Waffenbildschirm bietet dem Spieler die Wahl aus mehr als 250 Waffen, welche je eigene Unterfähigkeiten besitzen wie zum Beispiel das Beschwören von Schlangen, Angriffsmuster, Schwächen und Stärken besitzen.

Jede dieser Waffen kann einhändig oder zweihändig von der Einheit geführt werden. Wird eine Waffe einhändig geführt, kann die Einheit zwei unterschiedliche Waffen gleichzeitig tragen, was die Kombinationsmöglichkeiten zusätzlich erweitert.

Dazu kommen Sonderfähigkeiten von welchen die Einheit bis zu fünf erhalten kann. Diese Sonderfähigkeiten können alles von einem simplen Vampirbiss, welcher dem Gegner schadet und den eigenen Charakter heilt, bis zum Beschwören von Pfählen, die Gegner davon abhält sich zu wehren und zu bewegen, sein.

Im *Kleidungs*-Bildschirm werden dem Spieler mehr als 100 rein optische Möglichkeiten gegeben, die Einheiten visuell zu individualisieren. Für diese Individualisierung stehen dem Spieler fünf unterschiedliche Kategorien von Kleidung zur Verfügung; Kopf, Oberkörper, Beine, Hände und Schuhe.

Der darauf folgende Bildschirm ist der für diese wissenschaftliche Arbeit wichtigste. Der Statistikbildschirm, in welchem die unterschiedlichen Werte der Einheiten ganz genau modifiziert werden können.

Die Werte die in diesem Bildschirm angezeigt werden sind Lebenspunkte, Größe, Gewicht, Bewegungsgeschwindigkeit, Angriffsgeschwindigkeit und Schaden. Der Wert der Lebenspunkte ist hierbei der einzige, welcher mit einer genauen Zahl beschrieben werden kann. Die anderen Werte sind durch einen Multiplikator modifizierbar. Standardmäßig ist dieser Multiplikator 1.0. Für Gewicht, Größe und Bewegungsgeschwindigkeit sind dies standardisierte Werte und der Multiplikator wird immer den gleichen Effekt haben. Für Angriffsgeschwindigkeit und Schaden werden diese Multiplikatoren auf die Werte der gewählten Waffen angewandt, so kann eine Einheit mit einem Schadensmodifikator von zehn, aber einer simplen Faust immer noch weniger Schaden verursachen, als eine Einheit mit einer Kanone und einem

Schadensmodifikator von 0.5. Das gleiche zählt für die Angriffsgeschwindigkeit: eine Gatlinggun mit einem Geschwindigkeitsmodifikator von zehn kann eine Wand von Kugeln erzeugen während eine Muskete mit zehnfacher Angriffsgeschwindigkeit noch immer langsamer schießt als die Gatlinggun ohne Modifikation.

Auf dem finalen Bildschirm kann der Spieler dann die Einheit speichern. Hierzu kann dieser eine Beschreibung einfügen, ein Bild von der Einheit machen, mit welchem diese angezeigt wird und die Kosten der Einheit entweder automatisch erstellen lassen oder manuell beschreiben, ebenso wie dann final speichern falls die Einheit einen Namen besitzt.

Für die Zwecke dieser Arbeit wurde eine Einheit erstellt, welche von der Einheiten Basis her Stiffy ist und von der Stimme her ein Raptor mit maximaler Tonhöhe. Diese beiden Entscheidungen waren pur optisch und auditiv und beeinflussten den Verlauf der Tests nicht.

Das Bewegungsmuster wurde auf Nahkampfeinheit gestellt und das Zielverhalten auf Gegner. Für die Waffe wurde der zweihändig geführte Gehstock gewählt, da dieser eine recht hohe Trefferwahrscheinlichkeit besitzt und für die Zwecke der Reichweite modifiziert werden kann. Dazu kommt ein, für Totally Accurate Battle Simulator geringer Basisschaden, welcher es ermöglicht, dass Kämpfe nicht mit ein oder zwei Schlägen entschieden sind. Dadurch entstehen Kämpfe mit mehr Fokus auf den Werten und weniger auf der Chance den Gegner zu treffen.

Ein weiterer Vorteil des Gehstocks war es, dass dieser ein sehr weit verbreitetes Angriffsmuster hat, was für die Modifikation der Reichweite von immenser Bedeutung war. Ursprünglich war geplant, das Trainingsschwert zu verwenden, jedoch gibt es in der Schwertklasse kaum Unterschiede in der Reichweite, ohne besondere Effekte wie Feuerschaden dazu zu erhalten oder das Angriffsmuster komplett zu verändern. Die Alternative zum Gehstock waren noch die Lanzen, welche jedoch den Anspruch der hohen Trefferwahrscheinlichkeit nicht erfüllten wodurch ultimativ Weise der Gehstock gewählt wurde. Dieser gehört vom Angriffsmuster den stumpfen Waffen an, von welchen es eine Vielzahl in allen Größen und Formen gibt.

Für die Kleidung wurde der Basiseinheit eine simple schwarze Maske gegeben. Diese wurde gewählt, da jede Variante eine eindeutige Kopfbedeckung tragen sollte, um diese besser auseinander halten zu können. Dazu kamen simple Armreifen, welche die

Teamfarbe widerspiegeln, was förderlich für die Tests, in welchen gleiche Einheiten gegeneinander kämpfen, war.

Für jede Modifikation, abgesehen von der Reichweite, wurden stets die im Spiel vorhandenen Werteregler verwendet und durch die Möglichkeit der Eingabe von genauen Werten auf die der vorgesehenen Werte gesetzt. Diese Modifikation war die schnellste und einfachste der drei Spiele und konnte ohne Probleme abgeschlossen werden.

Eine Komplikation entstand jedoch bei der Suche nach der richtigen Reichweite. Hierbei wurde eine Vielzahl von Angriffsmustern analysiert und jede Waffe, die im Spiel vorhanden ist, anhand ihrer Reichweite und ihrem Angriffsmuster festgehalten. Finalerweise wurde dann der wie zuvor genannte Gehstock als Basis gewählt. Für die Modifikation der 75 prozentigen Reichweite wurde der simple Knüppel gewählt, für die 125 Prozent wurde der Baseballschläger Eins gewählt, für die Reichweite von 150 Prozent die Jarlsaxt und für die Reichweite von 200 Prozent das Guan.

Als Testumfeld wurde eine eigene Karte mithilfe des im Spiel implementierten Kartenerstellungswerkzeuges erstellt. Damit die Einheiten sich nicht von den anliegenden kämpfenden Einheiten beeinflussen lassen und damit die Einheiten keine Chance haben die Tests der anderen Subjekte zu manipulieren wie durch das Eingreifen oder das Ablenken der Einheiten, wurden breite Wände in großen Abständen voneinander platziert, um zehn ausreichend große Kampfarenen zu erschaffen. Hierbei wurde darauf geachtet, dass jegliche Einheiten ihre Waffen weiterhin ohne eine Behinderung schwingen können, da Behinderungen die Tests verfälschen würde, da nicht alle Einheiten die gleichen Grundbedingungen hätten.

## 4.2. Battle for Wesnoth Modifikation

Das Modifizieren des Spieles *The Battle for Wesnoth* erforderte, im Gegensatz zu *Totally Accurate Battle Simulator*, das Modifizieren des Codes des Spieles. Auch dies war sehr benutzerfreundlich, da *The Battle for Wesnoth* eine Open Source Anwendung mit einer Vielzahl an Foren ist.

Für die Modifikation des Spiels schlägt das offizielle Wiki vor, sich ein Beispiel an den im Spiel bereits vorhandenen Einheiten, Fraktionen, Animationen und ähnlichem zu nehmen. Hierzu war es nötig erst einmal herauszufinden, wo diese vom Spiel

gespeichert werden. Das offizielle Wiki verweist hierzu auf den Pfad  
 "C:/Users/<USER>/AppData/Local/Programs/Battle for Wesnoth <VERSION>/data".  
 Unter diesem Pfad sind jedoch bei einer Installation durch die Plattform Steam keine  
 Daten zu finden, da diese in einen steam-eigenen Ordner gespeichert werden. Das  
 kommt daher, dass Steam jegliche Hauptspeicherpfade von Spielen einheitlich im  
 Steamunterordner "T:/Steam/steamapps/common" abspeichert. Die genauen Daten  
 welche für diese Arbeit wichtig sind befinden sich unter  
 "T:/Steam/steamapps/common/wesnoth/data/core."

Der für das Modifizieren des Spiels relevante Code war dann wie folgend:

```

1  [ unit_type ]
2      id=Vampire Bat
3      name= _ "Vampire Bat"
4      race=bats
5      image="units/undead/bat-se-3.png"
6      profile="portraits/monsters/bat.png"
7      small_profile="portraits/monsters/bat.png~CROP
      (0,0,400,241)"
8      {TRAIT_FERAL_MUSTHAVE}
9      hitpoints=16
10     movement_type=smallfly
11     movement=8
12     experience=22
13     level=0
14     alignment=chaotic
15     advances_to=Blood Bat
16     cost=13
17     usage=scout
18     description= _ "Vampire bats are flying beasts that
      feed on the blood of other creatures. While their
      fangs are not very powerful, the health of victims
      is drained away along with their blood, and given
      to the Bats."
19     die_sound=bat-flapping.wav
20
21     [ resistance ]

```

```
22         cold=70
23     [ / resistance ]
24     [ defense ]
25         castle=50
26     [ / defense ]
27     [ ... ]
28     [ attack ]
29         name=fangs
30         description=_ "fangs"
31         icon=attacks/fangs-animal.png
32         type=blade
33         range=melee
34         damage=10
35         number=10
36         [ specials ]
37             {WEAPON_SPECIAL_DRAIN}
38         [ / specials ]
39     [ / attack ]
40     [ ... ]
41 [ / unit_type ]
```

Bei [...] folgend der [/defense] wurden 28 Zeilen an Animationen zwischen Angriffs und Verteidigungswerten ausgelassen, auf den [/attack] folgend wurden ebenso 18 weitere Zeilen für Angriffsanimationen ausgelassen.

Mit der Hilfe dieser Code Basis werden die Modifikationen des Spieles erschaffen. Der erste Schritt zur Modifikation dieser war es, die Spezialfähigkeit der Fledermaus, welche als Basis gewählt wurde zu entfernen, da diese die Lebenspunkte des Gegners der Einheit gutschreibt. Dies beeinflusst die Erfolgchancen, und würde die Tests verlängern. Der nächste Schritt war es die Nummer der Angriffe zu erhöhen, so dass die Angriffsgeschwindigkeit der Einheiten modifiziert werden kann.

Die Werte dieser Zeilen, die für die Arbeit modifiziert wurden sind, *hitpoints=100*, *damage=10* und *number=10* und für die Reichweite *[defense]castle=50*. Diese wurden je normalisiert auf durch zehn teilbare Werte geändert, um die prozentualen Veränderungen besser umsetzen zu können. So wurden die Lebenspunkte auf 100 gesetzt, der Schaden auf zehn und die Nummer der Angriffe ebenso auf zehn. Für

die Reichweite gab es in Battle for Wesnoth jedoch leider keinen eindeutigen modifizierbaren Wert. Dieser wurde daher durch die Trefferwahrscheinlichkeit ersetzt. Dieser Wert könnte jedoch sehr volatil sein, da die Trefferwahrscheinlichkeit anders funktioniert als die simple Reichweite der anderen Spiele.

Diese modifizierten Werte wurden dann für diese Bachelorarbeit mit den jeweiligen Werten multipliziert und getestet. Im Falle der *Hitpoints*, der Lebenspunkte zum Beispiel wurden die Werte auf 70, 110, 130 und 150 gesetzt.

Battle for Wesnoth sorgte bei der Erstellung der Modifikationen für überraschende Probleme, welche nicht in den Foren behandelt wurden. Das Problem war, dass jede Einheit eine einzigartige ID brauchte, was durch intensives Testen und Modifizieren aller Aspekte der Einheiten herausgefunden wurde. Dabei wurde die originale ID in der ersten modifizierten Einheit geändert, später jedoch nur Kopien mit jeweils nur geänderten Namen erstellt. Durch das Kopieren einer Einheit aus einer anderen Modifikation funktionierte dann auch diese nicht, obwohl deren ID einzigartig war. Nach der Überprüfung, ob es an fehlenden Art Assets lag und einen Tag an Fehlersuche wurden letztendlich die IDs einzeln modifiziert, was ultimativ den Fehler behob, wodurch das Testen der Einheiten begonnen werden konnte.

Um die Tests zu verbessern und zu verhindern, dass die Einheiten bei anderen Tests eingreifen wurden die Kosten der Einheiten auf eins gesetzt und die Bewegungsgeschwindigkeit auf null, da die Einheiten mit Hilfe der gezielten Kartenerstellung direkt nebeneinander gesetzt werden und sich nicht bewegen sollen.

Als Testumfeld wurde eine eigene Karte mithilfe des im Spiel implementierten Kartenerstellungswerkzeuges erstellt. Hierzu wurde die komplette Karte mit Feldern gefüllt, auf welchen man rekrutieren kann, und zwei Felder, auf welchen die Leiter der Armeen starten können und ebenfalls rekrutieren können. Für den Test können jegliche Felder je abwechselnd mit der Basiseinheit und mit der modifizierten Einheit befüllt und die Tests gestartet werden. So kann in jedem Testvorgang in der geringsten Zeit die maximale Anzahl an Einheiten getestet werden, da sich keine der Einheiten bewegen muss und diese einfach kämpfen.

Zusätzlich erhielten alle modifizierten Einheiten die Sonderfähigkeit Berserk, welche es diesen erlaubte bis zum Tod weiter zu kämpfen ohne die Testwerte zu verfälschen. Dies ersparte Zeit und ermöglichte, die Tests mit wenigen Aktionen innerhalb kürzester Zeit abzuschließen.



### 4.3. Rivals of Aether Modifikation

*Rivals of Aether* erfordert wie im Falle von Battle for Wesnoth das Modifizieren des Codes. Hierbei lag der Unterschied darin, dass in diesem Falle für jeden einzelnen modifizierten Charakter 19 Skripte individuell mit bis zu vier Einträgen modifiziert werden. Dies kommt daher, dass jede einzelne Attacke ihr eigenes Skript besitzt und die Einheiten selbst benannt und mit einer ID bestückt werden müssen. Dieses Modifizieren jedes einzelnen Skriptes erhöhte den Aufwand der Modifikationen immens.

Eine grundlegende Eigenheit welche für die Modifikation von Rivals of Aether betrachtet werden musste war, dass Kommazahlen im Code von Rivals of Aether nicht verwendbar sind, da dieses mit Integern, also Ganzzahlen, als Werten für die verwendeten Skripte arbeitet.

Die Skripte wurden wie bei Battle for Wesnoth von einer vom Offiziellen Wiki vorgeschlagenen simplen Einheit kopiert. Die modifizierten Skripte wurden dann unter dem Pfad "C:/Users/Tim/AppData/Local/RivalsofAether/workshop" innerhalb der Einheiten gespeichert.

Für die Modifikation der ersten Metrik, der des Schadens, wurden diese standardmäßig auf prozentual modifizierbare Zehnerschritten abgeädert. Da simples Aufrunden der Werte zu wenige Unterschiede der Angriffe erzeugen würden und den Unterschied zwischen leichten, mittleren und schweren Angriffen entfernen würde, wurden die Werte in Gruppen eingeteilt. Jeder Wert in der Spanne von eins bis fünf wurde hierbei auf zehn gesetzt, jeder Wert von sechs bis neun auf 20 und jeder Wert welcher größer oder gleich zehn wurde auf 30 gesetzt. Diese Einteilung erhält die Stärkenunterschiede der verschiedenen Fähigkeiten bei und ermöglicht es die Werte prozentual zu modifizieren.

Für die Modifikation der zweiten Metrik, der Reichweite, und die der vierten Metrik, der Angriffsgeschwindigkeit, wurden diese Werte für die prozentuale Modifizierbarkeit gerundet. Werte mit einer Endung von eins bis vier wurden hierbei abgerundet und Werte mit einer Endung von fünf bis neun wurden aufgerundet. Dies ergab prozentual modifizierbare Werte und verringerte die Varianz der Angriffe nur minimal. Werte mit geringen Reichweiten und Angriffsgeschwindigkeiten im Bereich zwischen 40 und 60 mit einer Endung von fünf und höher erhielten prozentual gesehen einen

sehr starken Bonus, während Werte mit einer Reichweite von mehr als 100 nur geringe prozentuale Unterschiede durch das Runden erhielten.

Im Falle der Lebenspunkte besaß Rivals of Aether keinen direkt modifizierbaren Wert. Die Alternative um diese zu testen wäre es gewesen, die Schadenswerte einer Einheit so zu modifizieren, dass diese simulieren wie die Siegesrate beeinflusst wäre, hätte die Einheit mehr Lebenspunkte. Dies wäre potentiell volatil und wenig aussagend und wurde aus diesem Grund für die Tests nicht berücksichtigt.

Wenn es durch die prozentuale Modifikation dazu kam, dass Werte mit Kommawerten entstehen wurde hierbei wieder, wie zuvor bei der Reichweite und der Angriffsgeschwindigkeit, gerundet.

Die gezielte Modifikation der Werte auf die prozentualen Modifikationen erfolgte mit der Hilfe eines Taschenrechners, um menschliche Fehler zu eliminieren. Bei jeder Modifikation wurde hierbei eine Liste auf einem Notizblock geführt auf welchem die vorhandenen Grundwerte und deren modifizierten Werte festgehalten wurden. Dies verringerte das Benötigen des Taschenrechners für jedes einzelne Skript und garantierte, dass keine Fehler durch falsche Erinnerungen geschehen.



## 5. Evaluation

In diesem Kapitel werden die Annahmen und modifizierten Anwendungen durch das praktische Testen mithilfe der in den Spielen vorhandenen künstlichen Intelligenz evaluiert.

Hierzu wird vorerst erläutert, welches Ziel diese Tests verfolgen, darauf folgend wird das Untersuchungsdesign konkretisiert und festgehalten. Das nächste Unterkapitel befasst sich dann mit der Durchführung der Tests und der Festhaltung der Ergebnisse. Diese werden abschließend ausgewertet und mit den zuvor getroffenen Annahmen verglichen.

### 5.1. Ziel

Das Ziel dieser Evaluation ist es, anhand der drei gewählten Genrebeispiele Rivals of Aether, Totally Accurate Battle Simulator und Battle for Wesnoth festzustellen, ob das Modifizieren von Metriken genreübergreifend angewendet werden kann.

### 5.2. Untersuchungsdesign

Die folgende Evaluation ist eine Forschung der quantitativen Art. Dies bedeutet, dass diese auf statistischen Auswertungen basiert. Weitergehend ist diese Evaluation um Repräsentativität bemüht und zielt in erster Linie auf Wahrscheinlichkeitsaussagen ab und die genreübergreifende Betrachtung dieser.

Die Auswertung der Daten wird mithilfe von Excel Tabellen vollzogen. Hierbei werden die Testversuche in Gruppen von zehn festgehalten und in die Tabelle eingetragen. Die Tabellen werden so eingerichtet, dass diese automatisch verrechnet werden. Diese Verrechnungen bieten die Möglichkeit auf einen Blick festzustellen, ob alle

zehn Tests vollständig durchgeführt wurden, wie hoch die Siegeswahrscheinlichkeit ist und wie viele Tests bereits durchgeführt wurden.

Diese Tabellen wurden für jedes Spiel erstellt und je in 16 Unterabteile eingeteilt. Je 70 beziehungsweise 75 Prozent, 110 beziehungsweise 125 Prozent, 130 beziehungsweise 150 Prozent und 150 beziehungsweise 200 Prozent. Und ebendiese vier Unterteilungen für jede der vier Metriken Schaden, Reichweite, Lebenspunkte und Angriffsgeschwindigkeit.

Für die Datensammlung wurden die in den Spielen vorhandenen künstlichen Intelligenzen genutzt. Dies verringerte den Testaufwand und ermöglichte es, diese Tests ohne die Zeit von Personen zu benötigen und ohne den Einfluss von wachsender Erfahrung der testenden Spieler in den Spielen einberechnen zu müssen. Dies erlaubt es die Tests innerhalb einer Woche zu beenden und zu jeder Tageszeit abzuschließen. Dennoch erfordert Rivals of Aether eine durchschnittliche Testzeit von drei bis sechs Minuten. Bei der großen Menge an benötigten Tests wird davon ausgegangen, dass diese tatsächlich jeden Tag der folgenden Woche kontinuierlich ausgeführt werden müssen, um es zu ermöglichen innerhalb der vorgesehenen Testphase fertig zu werden. Da die Tests mithilfe von künstlicher Intelligenz vollzogen werden, können die Tests von Totally Accurate Battle Simulator und Rivals of Aether potentiell komplett autonom nebeneinander laufen. Es wird bei beiden lediglich das Kommando gebraucht um die Tests zu starten und die Ergebnisse auszuzählen.

Zu den Testzeiten wird davon ausgegangen das die schnellsten die von Totally Accurate Battle Simulator sind. Hierbei wird davon ausgegangen, dass alle 30-120 Sekunden zehn Tests fertig sind. Dies kommt daher, dass die erstellte Testumgebung zehn Einheitenpaare auf einmal ermöglicht und dass diese nur wenig Zeit in Anspruch nehmen, da ein durchschnittlicher Kampf in Totally Accurate Battle Simulator rund 40 Sekunden mit den für die Tests modifizierten Grundeinheiten dauert.

Die längsten Tests werden die, wie zuvor angesprochenen, Rivals of Aether sein. Um sicher zu gehen, dass die Testergebnisse so genau wie möglich sind, werden die für das Spiel und Genre standardmäßigen drei Leben als Testumgebung gewählt. Hinzu kommt, dass jeder Test einzeln gestartet werden muss. Der Test muss also alle drei bis sechs Minuten bewertet und dann neu gestartet werden. Bei diesen Tests wird nur jeder vollständige Kampf als vollwertiger Test betrachtet.

Für das Testen werden die Bots in Rivals of Aether auf die höchstmögliche Stufe gestellt, um den Einfluss von Zufall durch schlechte künstliche Intelligenz, welche Fehler macht und zum Beispiel Selbstmord begeht, zu verringern und im Umkehrschluss den Einfluss der modifizierten Werte zu erhöhen.

Die Tests in Battle for Wesnoth benötigen ähnlich viel Input wie die in Rivals of Aether. Um zu garantieren, dass die Einheiten gegen die richtigen Gegnereinheiten kämpfen und die Testergebnisse nicht verfälschen, wurden die Tests der Einheiten jeweils manuell gestartet. Da die Einheiten jedoch die Kampffähigkeit Berserker erhalten haben kämpfen diese bis zum Tode ohne die Tests zu beeinflussen. Dies erspart das manuelle Beenden der Runde. Somit muss zwar jeder einzelne Testdurchgang einmal gestartet werden, kann jedoch mithilfe der speziell hierfür erstellten Karte ohne erneutes Laden jeweils mehr als 80 Tests ohne das erneute Platzieren der Einheiten durchführen. Durch die Speicherfunktion wurde bei Battle for Wesnoth dafür gesorgt, dass je 20 Einheiten pro Kategorie alle vier prozentualen Modifikationen von einer Metrik auf einmal getestet werden können. Nach dem diese einmal platziert wurden, wurde der aktuelle Stand gespeichert und der Test daraufhin fünf mal in Folge durchgeführt. Dies verringerte die, für das Platzieren der Einheiten, benötigte Zeit um 80 Prozent.

### **5.3. Durchführung**

Der erste Test wurde am 04. September 2023 durchgeführt. Mit diesem Datum fingen alle Tests an und wurden am 12. September 2023 beendet. Während dieser Zeitspanne wurden jeden Tag mehr als 8 Stunden getestet. Mehr als 70 Prozent der Testzeit fielen hierbei beim Spiel Rivals of Aether an. Für Totally Accurate Battle Simulator wurde mit der Erstellung der Einheiten und der Karte insgesamt weniger als 15 Stunden benötigt. Battle for Wesnoth benötigte rein zum Testen ebenso 15-20 Stunden und Rivals of Aether benötigte insgesamt rund 100 Stunden.

Für die Tests wurden die wie zuvor beschriebenen eigens erstellten Testumgebungen in Totally Accurate Battle Simulator und Battle for Wesnoth verwendet. Die Tests in Rivals of Aether wurden auf den Basisvarianten der im Spiel vorhandenen Karten veranstaltet. Diese Karten sind simple Karten mit je zwei bis vier Plattformen über der Basisplattform ohne Besondere, von der Karte ausgehende, Gefahren. Die nachfolgende Abbildung zeigt den Grundaufbau der drei Testumgebungen der Spiele.

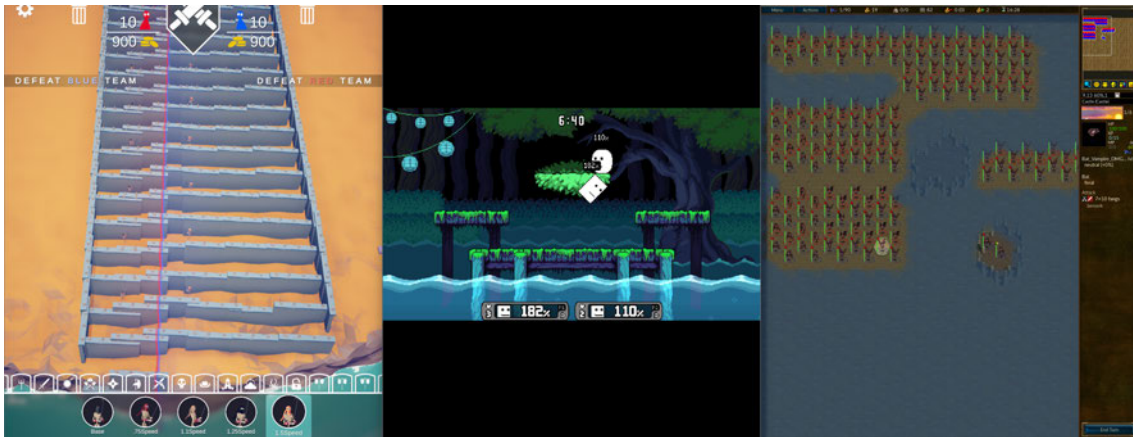


Abbildung 5.1.: Von links nach rechts: Totally Accurate Battle Simulator, Rivals of Aether und Battle for Wesnoth

Die Tests in Rivals of Aether wurden passiv im Hintergrund neben den anderen Tests ausgeführt. Rivals of Aether besitzt einen Fokus auf Controller-Support, weshalb diese Tests mithilfe eines Controllers gestartet werden konnten, auch wenn dessen Applikation nicht im Vordergrund ausgewählt ist. Dies ermöglichte es, die Tests der anderen Spiele gezielt mit der Maus zu steuern und nach jedem abgeschlossenen Rivals of Aether Test die Starttaste des Controllers zweimal zu betätigen, um die gewählten Charaktere und danach die Karte zu bestätigen. Um die Effizienz der Tests zu erhöhen und zu garantieren, dass diese direkt nach einander ablaufen, wurde am Anfang jedes Tests ein Timer auf fünf Minuten gestellt, sodass mithilfe des Alarms des Timers ein Hinweis erfolgte, dass der Test in wenigen Sekunden potentiell enden könnte. Damit wurde dafür gesorgt, dass in jeder aktiven Teststunde rund zwölf Tests durchgeführt werden konnten.

Während der ersten Tests von Battle for Wesnoth fielen einige Fehler auf, welche behoben werden mussten, um die Daten nicht zu verfälschen. Es kam zum Vorschein, dass die Einheiten zufällige Modifikatoren anhand ihrer Spezies erhalten und sich dies nicht mit der Zeile: `random_traits=no` entfernen lies, wie es der Fall gewesen wäre, wenn die Einheiten als direkte Einheit und nicht als Einheitentypus erstellt wurden. Der Vorteil des Erstellens des Einheitentypus war es, dass man diese in der Testumgebung mit der im Spiel vorhandenen Rekrutierfunktion ohne Probleme auf der Karte verteilen kann. Dieser Fehler der zufälligen Modifikatoren wurde behoben, indem die Spezies Fledermaus in den Kerndaten des Spiels überarbeitet wurde, und die Anzahl der zufälligen Modifikatoren auf null gesetzt wurde.

## 5.4. Auswertung

In diesem Unterkapitel werden die durch diese Arbeit gesammelten Daten ausgewertet und mit den zuvor im Kapitel 3.2 Gewichtungssannahmen getroffenen Annahmen verglichen. Die Daten werden quantitativ betrachtet und mit der Hilfe von Diagrammen anschaulich gemacht.

Die Spiele werden in diesem Kapitel erst einzeln in der Reihenfolge Totally Accurate Battle Simulator, Battle for Wesnoth und dann Rivals of Aether betrachtet. Anschließend werden die Kategorien der Spiele direkt miteinander verglichen.

In den nachfolgenden Diagrammen werden je die modifizierten Werte in einem dunklen Farbton dargestellt und sind oberhalb zu sehen, während die unmodifizierten Werte unterhalb der modifizierten Werte mit einer weniger kräftigen Farbe dargestellt werden. Hinzu kommt eine Anzeige, wie viele der Tests in Prozent von der unmodifizierten Einheit gewonnen wurden.

Zusätzlich werden die Werte der Tests im Bereich des Schadens stets mit Rot dargestellt, Werte der Tests im Bereich der Reichweite werden mit Blau markiert, Tests im Bereich der Lebenspunkte Grün und Angriffsgeschwindigkeitstests mit Gelb veranschaulicht.

Innerhalb der Diagramme ist die Reihenfolge immer von links nach rechts je 70 beziehungsweise 75 Prozent, 110 beziehungsweise 125 Prozent, 130 beziehungsweise 150 Prozent und 150 beziehungsweise 200 Prozent. Unterhalb der Säulen sind die jeweiligen Prozente angegeben.

Als Einschub vor den genauen Tests der Spiele im Folgenden die kompletten Daten der Tests der Einheiten mit identischen Metriken. Komplette Tabellen der drei anderen Spiele können im Anhang entnommen werden und die komplette PDF mit allen Tabellen kann im digitalen Abgabeordner gefunden werden. Die farblichen Markierungen entsprechen bei dieser Tabelle nicht dem Schaden, den Lebenspunkten und der Reichweite. Die farblichen Markierungen sind hierbei verwendet, um die unterschiedlichen Genres zu markieren.



## 5. EVALUATION

Baseline Tests									
	Totally Accurate Battle Simulator			Rivals of Aether			Battle for Wesnoth		
	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline
Sample 1-10	2	8	10	1	9	10	5	5	10
Sample 11-20	6	4	10	6	4	10	5	5	10
Sample 21-30	4	6	10	3	7	10	5	5	10
Sample 31-40	7	3	10	6	4	10	6	4	10
Sample 41-50	9	1	10	4	6	10	5	5	10
Sample 51-60	3	7	10	5	5	10	4	6	10
Sample 61-70	4	6	10	7	3	10	3	7	10
Sample 71-80	6	4	10	7	3	10	5	5	10
Sample 81-90	3	7	10	5	5	10	6	4	10
Sample 91-100	5	5	10	6	4	10	6	4	10
	49	51	100	50	50	100	50	50	100
	49,00%	51,00%	100,00%	50,00%	50,00%	100,00%	50,00%	50,00%	100,00%

Abbildung 5.2.: Die Tabellen der drei Spiele mit den Tests der Grundeinheiten

### 5.4.1. Totally Accurate Battle Simulator

Totally Accurate Battle Simulator war das erste Spiel bei dem die Tests abgeschlossen wurden. Die Werte dessen sind wie in den folgenden Diagrammen dargestellt ausgefallen.

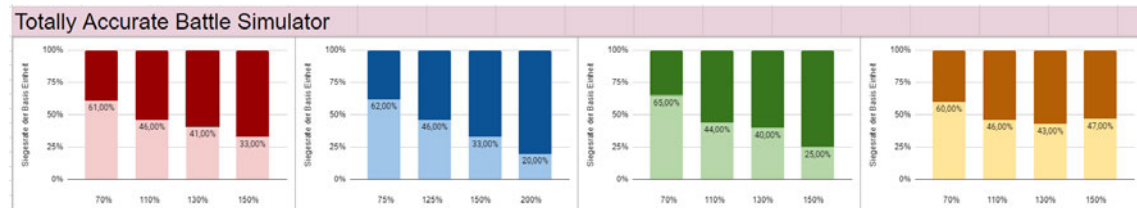


Abbildung 5.3.: Von links nach rechts: Schaden, Reichweite, Lebenspunkte und Angriffsgeschwindigkeit

Wie bereits zuvor erwähnt sind die modifizierten Werte der Tests die oberen Werte, welche in einer dunkleren Farbe markiert sind. Direkt auf den ersten Blick ist erkennbar, dass sich die Tendenz der Schadens-, Reichweitewerte und Lebenspunkte sehr vergleichbar und linear bis exponentiell entwickeln, während sich die Werte der Angriffsgeschwindigkeit sehr statisch verhalten.

#### Schaden

Zuerst werden die Werte der modifizierten Schadenswerte betrachtet. Die zuvor getroffenen Annahmen zur Schadensmetrik gingen davon aus, dass im Spiel Totally Accurate Battle Simulator die Einheit mit 70 Prozent des Schadens weniger als 43 Tests gewinnt, die Einheit mit 110 Prozent des Schadens ungefähr 55 der Tests

gewinnt, die Einheit mit 130 Prozent ungefähr 65 der Tests gewinnt und die Einheit mit den 150 Prozent des Schadens rund 70 der Tests gewinnt.

Die Tests ergaben, dass im Falle der 70 Prozent die Annahme richtig war; es wurden 39 Prozent der Tests von der 70 Prozent-Einheit gewonnen.

Bei den 110 Prozent war die Annahme ebenso sehr präzise, mit einer Siegesrate der modifizierten Einheit von 54 Prozent. Im Falle der 130 Prozent war die Annahme jedoch zu extrem. Die Tests ergaben eine Siegesrate von 59 Prozent. Dies kann damit erklärt werden, dass durch die Schadensvarianz in Totally Accurate Battle Simulator der Unterschied von 110 zu 130 Prozent nicht genug ist, um einen großen Breakpoint garantiert zu überschreiten, welcher die Siegesrate stark beeinflusst.

Die Tests mit der Einheit, welche eine Schadensmodifikation von 150 Prozent besitzt, waren im Gegensatz zu den 130 Prozent wieder sehr genau. Die Siegesrate betrug 67 Prozent, welche nur drei Prozent von den angenommenen 70 Prozent entfernt sind.

### **Reichweite**

Die zweite betrachtete Metrik der Reichweite fiel bei Totally Accurate Battle Simulator im Bereich der 75 Prozent und 125 Prozent fast identisch aus, wie die Tests der Schadensmodifikationen. Dies entsprach nicht den Annahmen, bei denen davon ausgegangen wurde, dass die Testwerte bei der Reichweite extremer ausfallen, als in den Schadenstests. Die Einheit mit 70 Prozent der Reichweite sollte so zum Beispiel nur ungefähr 35 Prozent der Tests gewinnen, erreichte jedoch eine Siegesrate von 38 Prozent. Dies könnte durch im Spiel vorhandene Varianz mit größeren Testmengen jedoch ausgeglichen werden.

Für den Fall der 125 Prozent war die finale Siegesrate identisch mit der Siegesrate der Schadensmodifikationen, nämlich einer Siegesrate von 54 Prozent. Auch hier wurde davon ausgegangen, dass die Werte extremere Maße annehmen und sich eher im Bereich einer Siegesrate von 60 Prozent befinden.

Bei den Tests mit einer Modifikation von 150 Prozent erreichte die modifizierte Einheit eine Siegesrate von 67 Prozent. Dies übertraf die Annahmen um zwei Prozent und lag damit im erwarteten Bereich.

Im Falle der 200 Prozentigen Reichweite lag die Siegesrate der Einheit bei ganzen 80 Prozent, was die Annahme um ganze 10 Prozent übertraf und bei den Tests für den einzigen Zehn zu Null Sieg innerhalb der Totally Accurate Battle Simulator Tests sorgte.

### **Leben**

Die Annahmen zur dritten Metrik, der Lebenspunkte, waren identisch mit denen der Schadensmetrik. Im Falle der Einheit mit 70 Prozent der Lebenspunkte erreichte die Einheit eine Siegesrate von 35 Prozent. Dies ist ganze acht Prozent schlechter als angenommen und unerwarteter Weise extremer, als die modifizierten Schadenswerte.

Im Test mit 110 Prozent der Lebenspunkte erreichte diese Einheit wie angenommen eine Siegesrate von 56 Prozent. Bei den Tests der Einheit mit 130 Prozent der Lebenspunkte erreichte die Einheit eine Siegesrate von 60 Prozent, was ebenso wie bei den Schadensmodifikationen geringer ist als erwartet und die Theorie der Breakpoints weiter bestärkt.

Mit 150 Prozent der Lebenspunkte erreicht die Einheit eine Siegesrate von 75 Prozent fünf Prozent mehr als erwartet und acht Prozent mehr als im Test der Schadenspunkte. Diese Tests lassen darauf schließen, dass im Spiel Totally Accurate Battle Simulator Lebenspunkte einen größeren Einfluss auf die Siegesrate einer Einheit haben als die Metrik des Schadens.

### **Angriffsgeschwindigkeit**

Die letzte betrachtete Metrik des Spiels Totally Accurate Battle Simulator ist die der Angriffsgeschwindigkeit. Die Betrachtung des Diagramms zeigt hierbei auf den ersten Blick, dass die Annahmen zum Einfluss der Angriffsgeschwindigkeit falsch lagen.

Im Fall mit 70 Prozent der Angriffsgeschwindigkeit erreichte die modifizierte Einheit eine 40 prozentige Siegesrate, fünf Prozent mehr als angenommen. Mit 110 Prozent der Geschwindigkeit erzielte die Einheit eine Siegesrate von 54 Prozent, sechs Prozent geringer als erwartet. Mit 130 Prozent der Geschwindigkeit erreichte die Einheit 57 Prozent Siegesrate, dies ist die höchste Siegesrate, die die Einheit mit der modifizierten Angriffsgeschwindigkeit erreicht hat. Mit 150 Prozent der Angriffsgeschwindigkeit erreichte sie gerade einmal 53 Prozent.

Die mit 130 und 150 Prozent modifizierten Einheiten erreichten beide einen Bruchteil der erwarteten Siegesrate. Für das Spiel *Totally Accurate Battle Simulator* kann also gesagt werden, dass das Modifizieren der Angriffsgeschwindigkeit das Schadenpotential einer Nahkampfeinheit erhöht, jedoch auf der anderen Seite durch das Erhöhen der Angriffsgeschwindigkeit auch das *ulkige* System weiter strapaziert wird und die Einheit zum Stolpern bringt, wodurch der Einfluss dieser nicht nur positiv ist sondern sogar auf erhöhten Werten die Volatilität erhöht und daher die Ergebnisse unberechenbarer macht.

### 5.4.2. Rivals of Aether

Die Tests des Spiels *Rivals of Aether* waren die letzten Tests, die beendet wurden, und wie in dem nachfolgenden Diagrammen dargestellt, wurden in diesem keine Tests zu den Lebenspunkten angestellt. Eine erste Betrachtung der Diagramme zeigt

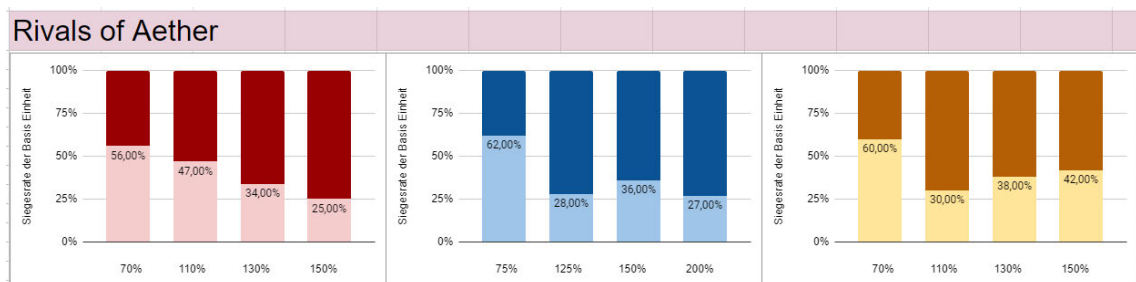


Abbildung 5.4.: Von links nach rechts: Schaden, Reichweite und Angriffsgeschwindigkeit

erneut einen linearen Anstieg der Siegeswahrscheinlichkeit bei dem modifizierten Schaden und stark variierende Werte bei der Angriffsreichweite und der Angriffsgeschwindigkeit.

#### Schaden

Zunächst werden die Werte der modifizierten Schadenswerte betrachtet. Wie bereits beim Spiel *Totally Accurate Battle Simulator* sind die zuvor getroffenen Annahmen zur Schadensmetrik wie folgt: die Einheit mit 70 Prozent des Schadens würde weniger als 43 Tests gewinnen, die Einheit mit 110 Prozent des Schadens ungefähr 55 der Tests, die Einheit mit 130 Prozent ungefähr 65 der Tests und die Einheit mit den 150 Prozent des Schadens würde rund 70 der Tests gewinnen.

Bei Rivals of Aether ergaben die Tests, dass im Falle der 70 Prozent die Annahme annähernd richtig war. Es wurden 44 Prozent der Tests von der 70 Prozent Einheit gewonnen. Die Annahmen zu den 110 Prozent waren, genau wie bei Totally Accurate Battle Simulator, sehr präzise mit einer Siegesrate der modifizierten Einheit von 53 Prozent.

Im Falle der 130 Prozent war die Annahme ebenso nah am erwarteten Ergebnis. Die Tests ergaben eine Siegesrate von 66 Prozent. Die Tests der Einheit mit einer Schadensmodifikation von 150 Prozent waren die ungenauesten im Vergleich zur Annahme, die Siegesrate betrug hierbei 75 Prozent, welche jedoch die angenommenen 70 Prozent auch nur um 5 Prozent übertraf.

### **Reichweite**

Bei der zweiten betrachteten Metrik, der der Reichweite, erreichte die Einheit mit 75 Prozent der Reichweite eine Siegesrate von 38 Prozent, was ungefähr der Annahme entspricht. Die Einheit mit 125 Prozent der Reichweite erreichte eine Siegesrate von 72 Prozent und übertraf damit alle Erwartungen. Die einzelnen Ergebnisse der Testgruppen zeigten auf, dass die Ergebnisse dieser Einheit sehr volatil und häufig Kämpfe mit einem Gleichstand der Siege sind. Es kann daher sein, dass sich die Siegesrate, mit einer größeren Testmenge, stark verändern würde.

Die Tests der Einheit mit 150 Prozent der Reichweite ergaben eine noch höhere Volatilität und so erreichte die Einheit nur eine 64 prozentige Siegesrate. Es schien beim Betrachten der Kämpfe so, als würde sich die Einheit selbst unter- und überschätzen. Es scheint als würde die Künstliche Intelligenz Probleme damit haben, den Effekt der stark erhöhten Reichweite anzuwenden. Mit 200 Prozent der Reichweite erreichte die modifizierte Einheit eine 73 prozentige Siegesrate. Auch hier erwies sich die Künstliche Intelligenz als weniger präzise im Vergleich zu den Tests mit 125 Prozent der Reichweite.

Zur Reichweite kann gesagt werden, dass der größte Einfluss auf die Siegesrate der Vorteil ist, die gegnerische Einheit zu treffen, bevor diese die Einheit selbst treffen kann. Den Einfluss, den diese Modifikation bringt, kann mit einer weiteren Erhöhung der Reichweite mit der in Rivals of Aether vorhandenen Künstlichen Intelligenz nicht weiter ausgebaut werden. Es scheint als gäbe es, wie bei den Schadenswerten von Totally Accurate Battle Simulator, einen Breakpoint bei der Reichweite von Rivals

of Aether, welcher beim Überschreiten dieser die Siegesraten immens verändert, während weitere Boni marginale Unterschiede der Siegesrate geben.

### **Angriffsgeschwindigkeit**

Bei der Angriffsgeschwindigkeit von Rivals of Aether bieten die steigenden Modifikationen einen überraschenden invertierten Einfluss. Im Gegensatz zu allen anderen Metriken besitzt die Einheit mit 110 Prozent der Geschwindigkeit die höchste Siegesrate mit 70 Prozent, welche mit der steigenden Modifikation stetig abfällt.

Mit 130 Prozent erreichte die modifizierte Einheit eine 62 prozentige Siegesrate. Die Einheit mit 150 Prozent der Angriffsgeschwindigkeit erreichte eine 58 Prozentige Siegesrate. Die Einheit mit 70 Prozent der Angriffsgeschwindigkeit erreichte eine Siegesrate von 40 Prozent.

Beim genauen Betrachten der Tests kann die Theorie der *verwirrten* Künstlichen Intelligenz, welche im Kapitel der Gewichtungsannahmen 3.2 getroffen wurde, als bestätigt gelten. Die Künstliche Intelligenz scheint mit jeder stärkeren Modifikation weniger Kontrolle zu besitzen.

Mit einer erhöhten Angriffsgeschwindigkeit kam auch eine verringerte Anzahl an aktiven Schaden verursachenden Frames. Die Einheiten schienen mit erhöhten Modifikationen zwar mehr Angriffe auszuführen, machten jedoch auch mehr Fehler und schafften es seltener die Angriffe der Unmodifizierten zu parieren, während ebendiese unmodifizierte Einheit die Angriffe der Einheiten mit hoher Präzision parierte.

### **5.4.3. Battle for Wesnoth**

*Battle for Wesnoth* ist das Spiel mit der geringsten beziehungsweise mit der höchsten Berechenbarkeit an Varianz. Es gibt in *Battle for Wesnoth* eine Funktion um im Vorhinein zu sehen, wie hoch die Chance einer Einheit zu gewinnen ist und mit wie vielen Lebenspunkten diese gewinnen wird. Diese Berechnungen des Spiels sind fast in allen Fällen eins zu eins in den finalen Tests widergespiegelt.

### **Schaden**

Die erste Metrik, welche betrachtet wurde, ist wieder die der Schadensmodifikation. Direkt die erste Einheit, welche mit 70 Prozent des Schadens modifiziert ist, erreicht gerade einmal eine neun prozentige Siegesrate. Dies ist wesentlich weniger,

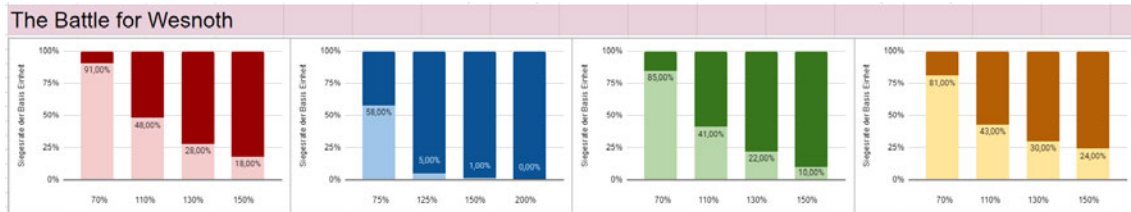


Abbildung 5.5.: Von links nach rechts: Schaden, Reichweite, Lebenspunkte und Angriffsgeschwindigkeit

als die zuvor angenommenen rund 43 Prozent. Im Fall der 110 Prozent erreichte die modifizierte Einheit dann eine Erfolgsrate von 52 Prozent, nur 3 Prozent weniger als angenommen. Im Fall der 130 Prozent erreichte die modifizierte Einheit eine Siegesrate von 72 Prozent und übertraf damit alle Annahmen. Mit 150 Prozent des Schadens erreichte die Einheit eine Siegesrate von 82 Prozent und wies auf eine potentiell exponentielle Verbindung zwischen dem Wachstum des modifizierten Wertes und der Siegesrate hin.

### Reichweite

Bei der zweiten Metrik der Reichweite waren die Daten, wie zuvor in der Annahme befürchtet, extrem volatil und wenig aussagekräftig. Im Falle der 75 prozentigen "Reichweite" erzielte die Einheit wider Erwartens eine Siegesrate von 42 Prozent.

Im Falle der 125 prozentigen Reichweite erzielte die Einheit eine Siegesrate von 96 Prozent. Bei einer Reichweite von 150 Prozent erzielte die Einheit eine Siegesrate von 99 Prozent. Bei 200 Prozent der Reichweite erreichte die modifizierte Einheit dann eine Rate von 100 Prozent.

### Lebenspunkte

Die dritte getestete Metrik, die der Lebenspunkte, fiel bei Battle for Wesnoth vergleichbar mit denen der modifizierten Schadenswerte aus jedoch wieder, wie schon bei Totally Accurate Battle Simulator, mit stärkerem Effekt als der Schaden. Die Einheit mit 70 Prozent der Lebenspunkte erzielte eine Siegesrate von 15 Prozent. Dies ist sechs Prozent mehr als die Einheit mit modifiziertem Schaden und fast 30 Prozent schlechter als angenommen.

Die Einheit mit 110 Prozent der Lebenspunkte erreichte hingegen eine Siegesrate von 59 Prozent. Dies liegt noch fast im angenommenen Bereich, mit nur vier Prozent mehr als vorhergesagt.

Die Einheit mit 130 Prozent der Lebenspunkte erreichte eine Siegesrate von 78 Prozent und damit 13 Prozent mehr als angenommen, während die Einheit mit 150 Prozent der Lebenspunkte eine 90 prozentige Siegesrate erreichte, ganze 20 Prozent mehr als angenommen.

### **Angriffsgeschwindigkeit**

Die Angriffsgeschwindigkeit sorgte in dieser Testumgebung für schwächere Ergebnisse als die Metriken der Schadens- und Lebenspunkte. Dies widersprach den Annahmen, die davon ausgingen, dass die Siegesrate der Angriffsgeschwindigkeit fast identisch zu der Siegesrate der modifizierten Lebenspunkte ist.

Die Einheit mit 70 Prozent der Angriffsgeschwindigkeit erreichte eine Siegesrate von 19 Prozent, 20 Prozent weniger als angenommen. Die Einheit mit 110 Prozent der Angriffsgeschwindigkeit erreichte eine Siegesrate von 57 Prozent, welche ungefähr im angenommenen Bereich liegt.

Mit 130 Prozent der Angriffsgeschwindigkeit erreichte die Einheit eine 70 prozentige Siegesrate, welche nur fünf Prozent über der angenommenen Siegesrate liegt. Die Einheit mit 150 Prozent der Angriffsgeschwindigkeit erreichte eine Siegesrate von 76 Prozent, auch diese lag im Bereich der angenommenen 70 Prozent.

### **5.4.4. Vergleich der Genre**

Anschließend werden die Daten der verschiedenen Genre untereinander anhand der in den Tests festgehaltenen Daten verglichen.

#### **Schaden**

Die erste Metrik ist die des modifizierten Schadens. Anhand des nachfolgenden Diagramms 5.6 ist erkennbar, dass die Annahmen der exponentiellen Effekte sehr akkurat waren. Vorallem im Fall von Battle for Wesnoth kann ein exponentieller Effekt der Boni festgestellt werden.



Im Bereich der 110 prozentigen Modifikation kann festgestellt werden, dass die Effekte der Modifikationen genreübergreifend sehr vergleichbar sind, mit einer Varianz zwischen den Spielen von ein bis zwei Prozent.

Die Effekte der 70 Prozent in den Spielen Totally Accurate Battle Simulator und Rivals of Aether sind ebenso weiterhin vergleichbar mit einer Varianz von 5 Prozent im Gegensatz zur Siegesrate des Spiels Battle for Wesnoth, welche eine Differenz von bis zu 35 Prozent besitzt.

Im Falle der 70 prozentigen Modifikation besteht die größte Differenz zwischen den Genres in der Metrik der Schadensmodifikation. Ab einer Modifikation von mehr als 130 Prozent sind die Unterschiede zwischen den Genres zu gravierend um sagen zu können, das Balancing genreübergreifend angewandt werden kann.

Bei 130 Prozent besteht eine Varianz von sechs bis 13 Prozent, während bei der Modifikation von 150 Prozent eine Varianz von sieben bis 15 Prozent besteht. Ab 130 Prozent ist die Varianz zwischen Battle for Wesnoth und Rivals of Aether tatsächlich vergleichbar mit der Varianz zwischen Rivals of Aether und Totally Accurate Battle Simulator. Rivals of Aether bildet ab der Modifikation von 110 Prozent bei jeder weiteren Modifikation fast die exakte Mitte zwischen den beiden anderen Spielen.

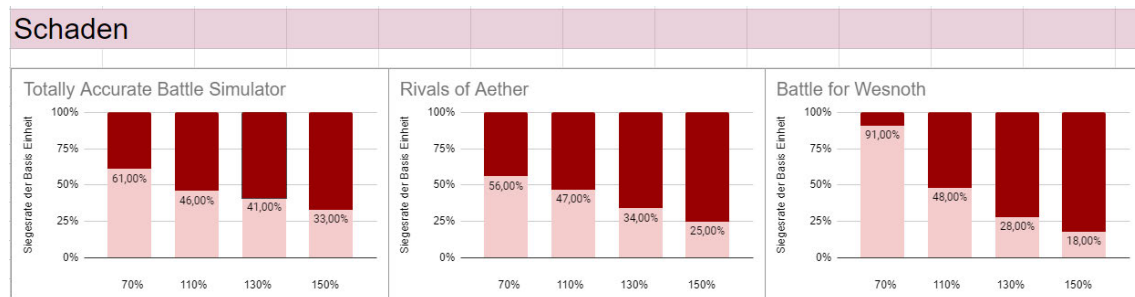


Abbildung 5.6.: Balkendiagramme der Schadens Metrik aller Genre

### Reichweite

Bei der zweiten Metrik, der Reichweite, bestehen im Bereich der 70 prozentigen Modifikation diesmal vergleichbare Werte.

Die Effekte der 70 prozentigen Reichweite sind in den Spielen Rivals of Aether und Totally Accurate Battle Simulator identisch und die Varianz zu Battle for Wesnoth beträgt ebenso nur vier Prozent.

Die Modifikation der 125 Prozent erreichte eine Varianz von 18 bis 41 Prozent und ist daher nicht genreübergreifend anwendbar. Ab der Modifikation der 125 Prozent sind die Werte des Spiels Battle for Wesnoth so extrem, dass diese für die folgenden Vergleiche nicht mehr genannt werden.

Die Modifikation der 150 Prozent hingegen erreichte bei den Spielen Rivals of Aether und Totally Accurate Battle Simulator vergleichbare Werte von 64 und 67 Prozent. Bei der Modifikation der 200 prozentigen Reichweite verhält sich die Siegesrate von Rivals of Aether stärker steigend als die von Totally Accurate Battle Simulator, die Werte sind daher nicht mehr genreübergreifend vergleichbar.

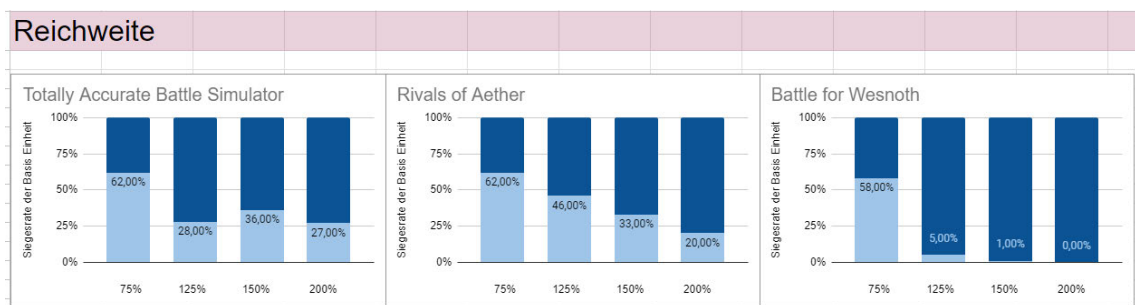


Abbildung 5.7.: Balkendiagramme der Reichweiten Metrik aller Genre

### Lebenspunkte

Für die Metrik der Lebenspunkte konnten nur die Spiele Battle for Wesnoth und Totally Accurate Battle Simulator verglichen werden. Anhand des Diagramms 5.8 sind große Differenzen in allen Metriken außer der Modifikation der 110 Prozent erkennbar.

Der gravierendste Unterschied ist wieder erkennbar an der Modifikation der 70 Prozent. Die Differenz zwischen den Siegesraten liegt bei ganzen 20 Prozent. Die Metrik der 110 Prozent bietet vergleichbare Werte von 56 und 59 Prozent, welche die Theorie des genreübergreifenden Balancieren für geringe Modifikationen unterstützen.

Die Metrik der 130 Prozent deutet bei Totally accurate Battle Simulator darauf hin, dass diese keinen neuen Breakpoint erreichte und daher fast identisch mit der Modifikation der 110 Prozent ist. Die Differenz zwischen den beiden Spielen ist hierbei wieder sehr groß mit ganzen 18 Prozent, ein weiterer Beweis, wie wichtig es ist beim Modifizieren der Werte auf die im Spiel vorhandenen Breakpoints zu achten.

Der Unterschied der Werte bei der Modifikation der 150 Prozent sind wieder näher bei einander mit 15 Prozent Differenz. Trotz der Anomalie des fehlenden Breakpoints im Spiel Totally Accurate Battle Simulator bei 130 Prozent ist anhand der 150 prozentigen Modifikation der stärkere exponentielle Effekt der Modifikationen im Spiel Battle for Wesnoth erkennbar.

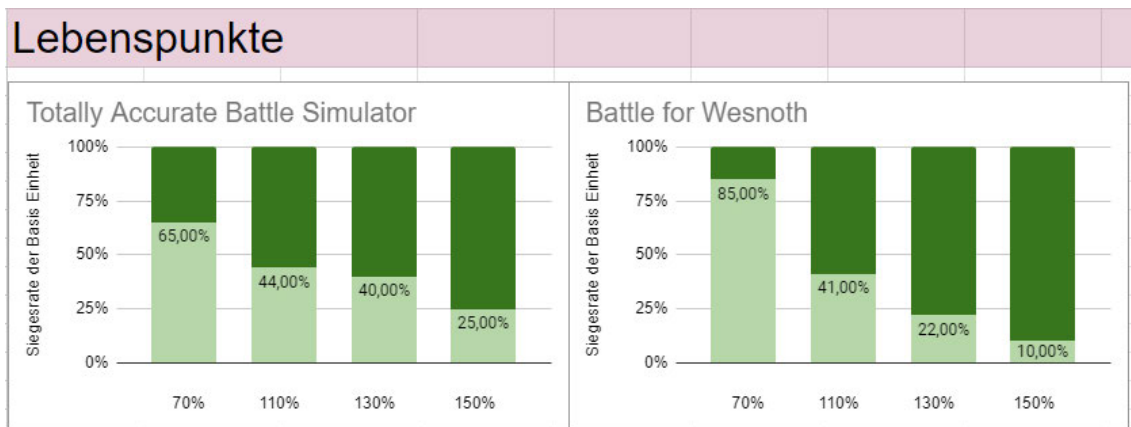


Abbildung 5.8.: Balkendiagramme der Lebenspunkte Metrik von Totally Accurate Battle Simulator und Battle for Wesnoth

### Angriffsgeschwindigkeit

Die finale Metrik der Tests ist die der Angriffsgeschwindigkeit. Bei dieser lässt sich bei den Spielen Totally Accurate Battle Simulator und Rivals of Aether wieder ein Trend erkennen. Die Angriffsgeschwindigkeit sorgte bei beiden Spielen weder für einen linearen noch einen exponentiellen Anstieg der Siegesrate mit dem Anstieg der Modifikation, sondern tatsächlich eher ein stagnanter Verlauf der Siegesrate. Selbst bei Battle for Wesnoth ist nur ein linearer Anstieg der Siegesrate erkennbar.

Bei der 70 prozentigen Modifikation sind die Siegesraten bei Totally Accurate Battle Simulator und Rivals of Aether erneut identisch, während der Effekt bei Battle for Wesnoth 21 Prozent stärker ist als bei den anderen Spielen.

Bei der Modifikation der 110 Prozent ist eine Vergleichbarkeit zwischen Totally Accurate Battle Simulator und Battle for Wesnoth erkennbar. Die Varianz zwischen den beiden Spielen liegt bei nur drei Prozent, während die Varianz zu Rivals of Aether bis zu 16 Prozent beträgt.

Bei 130 Prozent tritt das erste Mal der Trend der Stagnation auf. Totally Accurate Battle Simulator und Rivals of Aether befinden sich bei einer Angriffsrate von 130

Prozent beide bei je 57 und 62 Prozent mit einer Differenz von fünf Prozent zwischen ihnen. Während die modifizierte Einheit des Spiels Totally Accurate Battle Simulator drei Prozent der Kämpfe weniger gewinnt, verlor die Einheit in Rivals of Aether ganze acht Prozent der Kämpfe mehr. Die ansteigende Angriffsgeschwindigkeit schwächte die Einheit mit der 130 prozentigen Modifikation. Battle for Wesnoth erreichte eine Siegesrate von 70 Prozent und damit eine Varrianz von acht bis 13 Prozent im Vergleich zu den anderen Spielen.

Bei der Modifikation mit 150 Prozent der Geschwindigkeit verloren beide Testsubjekte Rivals of Aether und Totally accurate Battle Simulator weiterhin mehr Kämpfe als bei den Tests der 130 Prozent. Battle for Wesnoth hingegen erlangte auch nur 6 Prozent mehr Siege als im vorherigen Test. Die Stagnation kann bei der Angriffsgeschwindigkeit in allen drei Genren festgestellt werden.

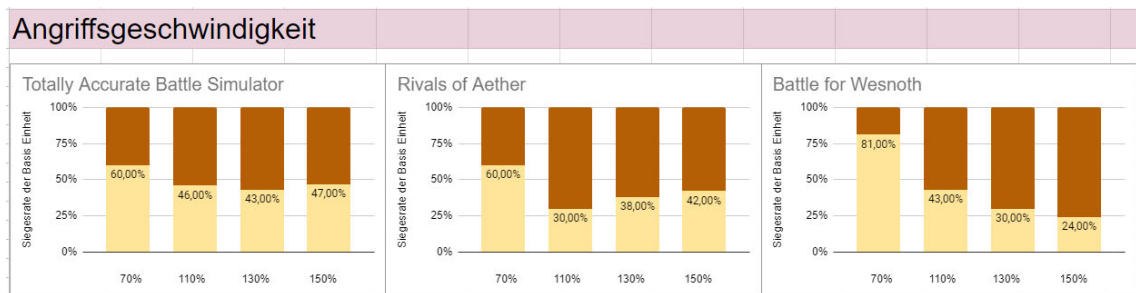


Abbildung 5.9.: Balkendiagramme der Angriffsgeschwindigkeits Metrik aller Genre



## 6. Schlussbetrachtung

Im finalen Kapitel dieser wissenschaftlichen Arbeit werden im ersten Unterkapitel die Funde dieser Arbeit zusammengefasst und noch einmal konkret aufgelistet. Abschließend wird ein Ausblick aufgestellt der beschreibt, welchen Effekt die Funde auf das Balancing in der Videospelindustrie haben können.

### 6.1. Zusammenfassung

Die Tests dieser Bachelorarbeit ergaben, dass Modifikationen von zehn Prozent in den Metriken, Schadenspunkte und Lebenspunkte genreübergreifend vergleichbare Werte generieren.

Hierbei ist jedoch darauf zu achten, ob diese Modifikationen mit den in den Spielen vorhandenen Breakpoints interagieren. Überschreitet die Modifikation in einem Spiel einen Breakpoint und befindet sich im anderen innerhalb dessen, sind die Unterschiede der Modifikationen genreübergreifend immens. Überschreiten beide Modifikationen die gleiche Menge an Breakpoints sind die Änderungen des Balancings jedoch wieder vergleichbar.

Ein weiterer Aspekt, der sich für das genreübergreifende Balancing als markant dargestellt hat, ist der Aspekt des exponentiellen Effekts einer Metrik auf die Siegesrate innerhalb eines Spiels. Mit steigender Modifikation der Werte erhöht sich der Effekt dieser Modifikationen zum Beispiel in Battle for Wesnoth wesentlich stärker als die Siegesrate der anderen Spiele.

Diese Anstiege treten jedoch erst bei den Modifikationen der 130 Prozent oder der 70 Prozent auf. Der anfängliche langsame Anstieg weist hierbei auf eine exponentielle Kurve hin und sorgt dafür, dass geringe Modifikationen genreübergreifend fast

identisch sind, jedoch aber größere Modifikationen mit Werten von mehr als 30 Prozent, je nach Genredifferenz immense balancingtechnische Unterschiede offenbaren können.

Ein weiterer Aspekt, der bei den Tests heraustrat, ist, dass sich Spiele wie Rivals of Aether mit künstlicher Intelligenz und Totally Accurate Battle Simulator mit einer hohen Genrevarianz anders verhalten und besser miteinander vergleichbar sind, als das Spiel Battle for Wesnoth, welches eine geringe und berechenbare Varianz durch die genau angegebenen Trefferwahrscheinlichkeiten bietet. Die varianzreichen Spiele Rivals of Aether und Totally Accurate Battle Simulator erzielten bei den Tests oft ähnliche Werte mit geringer Varianz auch außerhalb der zehn Prozent Modifikation. Es kann davon ausgegangen werden, dass ein weiteres varianzarmes Spiel und Genre ähnliche Ergebnisse wie Battle for Wesnoth produzieren würde.

Weitergehend hat sich die Trefferwahrscheinlichkeit in Battle for Wesnoth als komplett unvergleichbar mit der Angriffsreichweite der anderen Spiele herausgestellt. Die Vergleichbarkeit der genreübergreifenden Metriken stellt sich auch hier als immens wichtiger Aspekt heraus.

Es kann also abschließend gesagt werden, dass geringe Modifikationen auch genreübergreifend angewandt werden können, jedoch nicht ohne genau zu betrachten welche zusätzlichen Faktoren innerhalb dieser aktiv sind. Diese Bachelorarbeit und die in dieser durchgeführten Tests haben ergeben, dass beim genreübergreifenden Balancing vier Kernfaktoren beachtet werden müssen.

Der erste Faktor ist das Betrachten der Genrevarianz. Es ist wichtig, die zu vergleichenden Genre in ihrer Varianz zu betrachten. Besitzt eines der Genre weniger Aspekte, die einen schwer berechenbaren Einfluss auf das Geschehen haben, als das andere sind grobe Modifikationen mit mehr als 30 prozentigen Änderungen schwer möglich, wie an der Differenz zwischen den Testwerten der Metrik Angriffsgeschwindigkeit von Rivals of Aether und Totally accurate Battle Simulator im Vergleich zu den Werten von Battle for Wesnoth erkennbar ist.

Der zweite zu betrachtende Faktor ist der der Breakpoints. Breakpoints sind Punkte, die in einer varianzarmen Testumgebung entscheiden, wie viele Aktionen benötigt werden, um einen Sieg zu erlangen. Dies wird im Beispiel des dritten Kapitels deutlich: Besitzt eine Einheit 100 Lebenspunkte ist es für diese unwichtig, ob der

Angreifer 20 oder 24 Schaden verursacht. Ist diese Zahl nicht variabel braucht der Angreifer mit jedem dieser Werte weiterhin fünf Schläge.

Dieser Faktor der Breakpoints wurde vor allem bei den Tests der Lebenspunkte im Spiel *Totally Accurate Battle Simulator* festgestellt. Der Unterschied einer 110 prozentigen Modifikation im Vergleich zu einer 130 prozentigen Modifikation betrug gerade einmal vier Prozent, während der Unterschied von 130 Prozent zu 150 Prozent der Lebenspunkte ganze 15 Prozent Differenz erwies.

Das unterschiedlich starke exponentielle Wachstum des Effekts, den eine Metrik in einem Spiel oder Genre besitzt, ist der dritte markante Faktor, den diese Arbeit herausstellte. Dieser stellte sich bei den Tests vor allem im Spiel *Battle for Wesnoth* heraus.

Die Werte von *Battle for Wesnoth* waren in fast jeder Metrik extremer als die Werte der anderen beiden Spiele. Dies ist besonders erkennbar bei der Metrik der Schadenspunkte, bei welcher im Bereich der 110 prozentigen Modifikation fast identisch waren, jedoch im Bereich der 70 Prozent und 150 Prozent immense Unterschiede baten.

Der letzte und markanteste Faktor, ist der der Vergleichbarkeit der genreübergreifenden Metriken. Die Metrik der Reichweite wurde für *Battle for Wesnoth* mithilfe der Trefferwahrscheinlichkeit dargestellt. Dies sorgte für Werte, welche sehr stark von den Testwerten der anderen beiden getesteten Spiele variierten. Was diese Metrik als nicht genreübergreifend vergleichbar darstellte.

Zur Wahl der Genre und Spiele kann abschließend gesagt werden, dass *Rivals of Aether* zwar immensen Aufwand mit sich brachte und ein Großteil der zeitlichen Ressourcen für die Tests und Modifikationen innerhalb dieses Spiels verbraucht wurden, aber dennoch als Genre eine solch große und relevante Abwechslung zu den anderen beiden Genres bot, dass bei einer Weiterführung dieser Tests *Rivals of Aether* wieder auftreten würde.



## 6.2. Ausblick

Abschließend kann gesagt werden, dass die in der Bachelorarbeit herausgestellten Faktoren beim genreübergreifenden Balancing stets beachtet werden müssen.

Kann der Spieleentwickler sicherstellen, dass diese Faktoren für beide Genre geringe Varianz bieten, kann dieser genreübergreifend geringe Balancingänderungen und Erkenntnisse übernehmen. Falls zwischen den beiden Genres kaum eine Differenz der genannten Faktoren besteht, können größere Balancingänderungen mindestens als Tendenz betrachtet angewendet werden, wie an den Spielen *Totally Accurate Battle Simulator* und *Rivals of Aether* in der Metrik der Angriffsgeschwindigkeit sichtbar ist.

## Literaturverzeichnis

- [Ada13] Ernest Adams: *Fundamentals of Game Design, Third Edition*, Bd. 3, New Riders, 2013, ISBN 9780133435726.
- [Ars09] Dominic Arsenault: *Video Game Genre, Evolution and Innovation, Eudamos. Journal for Computer Game Culture*, Bd. 3(2):S. 149–176, 2009.
- [BAA<sup>+</sup>16] Marlene Beyer, Aleksandr Agureikin, Alexander Anokhin, Christoph Laenger, Felix Nolte, Jonas Winterberg, Marcel Renka, Martin Rieger, Nicolas Pflanzl, Mike Preuss und Vanessa Volz: *An integrated process for game balancing*, in *2016 IEEE Conference on Computational Intelligence and Games (CIG)*, S. 1–8, 2016.
- [BG20] Alexander Becker und Daniel Görlich: *What is Game Balancing? An Examination of Concepts, Paradigm Plus*, Bd. 1(1):S. 22 – 41, 2020.
- [Gam] Landfall Games: *Totally Accurate Battle Simulator*, URL: <https://landfall.se/totally-accurate-battle-simulator>, Aufgerufen am 16.07.2023.
- [Gar16] Chaim Gartenberg: *Totally Accurate Battle Simulator brings the wacky joy of fighting action figures to life / War never changes*, 2016, URL: <https://www.theverge.com/2016/12/4/13822182/totally-accurate-battle-simulator-landfall-games-pc-wacky>, Aufgerufen am 19.07.2023.
- [Jaf] Alexander Benjamin Jaffe: *Understanding Game Balance with Quantitative Methods*, Diplomarbeit, University of Washington.
- [Ket] Miikka Ketonen: *Designing a 2D fighting game*, Diplomarbeit, Kajaanin Ammattikorkeakoulu University of Applied Sciences.
- [Sch81] Thomas Schatz: *Formulas Filmmaking and the Studio System*, Random House, 1981, ISBN 9780075536239.

- [Stua] Aether Studios: *Press Kit*, URL: <https://rivalsofaether.com/press/>, Aufgerufen am 16.07.2023.
- [Stub] Aether Studios: *Rivals of Aether*, URL: [https://store.steampowered.com/app/383980/Rivals\\_of\\_Aether/](https://store.steampowered.com/app/383980/Rivals_of_Aether/), Aufgerufen am 16.07.2023.
- [Stuc] Aether Studios: *What is Rivals of Aether?*, URL: <https://rivalsofaether.com/>, Aufgerufen am 16.07.2023.
- [Wei] Jenny Weissbach: *Identifikation eines Spielertyps in dem Soulslike-Genre*, Diplomarbeit, Hochschule Mittweida, University of Applied Science.
- [RCL] Álvaro Gutiérrez Rodríguez, Carlos Cotta und Antonio J. Fernández Leiva: *An Evolutionary Approach to Metroidvania Videogame Design*, *Conferencia de la Asociación Española para la Inteligencia Artificial*.

# Anhang



# A. Analysedokumente

## A.1. Exceltabellen

Totally accurate Battle Simulator																																																							
Damage														Range																																									
Baseline VS 70%							Baseline VS 110%							Baseline VS 130%							Baseline VS 150%							Baseline VS 75%							Baseline VS 125%							Baseline VS 150%							Baseline VS 200%						
Sample	1-10	11-20	21-30	31-40	41-50	51-60	Sample	1-10	11-20	21-30	31-40	41-50	51-60	Sample	1-10	11-20	21-30	31-40	41-50	51-60	Sample	1-10	11-20	21-30	31-40	41-50	51-60	Sample	1-10	11-20	21-30	31-40	41-50	51-60	Sample	1-10	11-20	21-30	31-40	41-50	51-60	Sample	1-10	11-20	21-30	31-40	41-50	51-60	Sample	1-10	11-20	21-30	31-40	41-50	51-60
61.0%	39.0%	100.0%	46.0%	54.0%	100.0%	41.0%	100.0%	41.0%	59.0%	100.0%	33.0%	67.0%	100.0%	62.0%	38.0%	100.0%	46.0%	54.0%	100.0%	33.0%	100.0%	46.0%	54.0%	100.0%	33.0%	67.0%	100.0%	62.0%	38.0%	100.0%	46.0%	54.0%	100.0%	33.0%	100.0%	46.0%	54.0%	100.0%	33.0%	67.0%	100.0%	62.0%	38.0%	100.0%	46.0%	54.0%	100.0%	33.0%							
Average Hit Percentage														Average Hit Percentage																																									
191.00														191.00																																									
45.25%														45.25%																																									
218.00														218.00																																									
400.00														400.00																																									
HP														Speed																																									
Baseline VS 70%							Baseline VS 110%							Baseline VS 130%							Baseline VS 150%							Baseline VS 70%							Baseline VS 110%							Baseline VS 130%							Baseline VS 150%						
Sample	1-10	11-20	21-30	31-40	41-50	51-60	Sample	1-10	11-20	21-30	31-40	41-50	51-60	Sample	1-10	11-20	21-30	31-40	41-50	51-60	Sample	1-10	11-20	21-30	31-40	41-50	51-60	Sample	1-10	11-20	21-30	31-40	41-50	51-60	Sample	1-10	11-20	21-30	31-40	41-50	51-60	Sample	1-10	11-20	21-30	31-40	41-50	51-60	Sample	1-10	11-20	21-30	31-40	41-50	51-60
61.0%	39.0%	100.0%	46.0%	54.0%	100.0%	41.0%	100.0%	41.0%	59.0%	100.0%	33.0%	67.0%	100.0%	62.0%	38.0%	100.0%	46.0%	54.0%	100.0%	33.0%	100.0%	46.0%	54.0%	100.0%	33.0%	67.0%	100.0%	62.0%	38.0%	100.0%	46.0%	54.0%	100.0%	33.0%	100.0%	46.0%	54.0%	100.0%	33.0%	67.0%	100.0%	62.0%	38.0%	100.0%	46.0%	54.0%	100.0%	33.0%							
Average Hit Percentage														Average Hit Percentage																																									
191.00														191.00																																									
45.25%														45.25%																																									
218.00														218.00																																									
400.00														400.00																																									
Global Average														Global Average																																									
191.00														191.00																																									
45.25%														45.25%																																									
218.00														218.00																																									
400.00														400.00																																									
extremely volatile														extremely volatile																																									

Abbildung A.1.1.: Gesamtübersicht der Tests im Spiel Totally Accurate Battle Simulator

Rivals of Aether																																																							
Damage														Range																																									
Baseline VS 70%							Baseline VS 110%							Baseline VS 130%							Baseline VS 150%							Baseline VS 75%							Baseline VS 125%							Baseline VS 150%							Baseline VS 200%						
Sample	1-10	11-20	21-30	31-40	41-50	51-60 <th>Sample</th> <th>1-10</th> <th>11-20</th> <th>21-30</th> <th>31-40</th> <th>41-50</th> <th>51-60 <th>Sample</th><th>1-10</th><th>11-20</th><th>21-30</th><th>31-40</th><th>41-50</th><th>51-60 <th>Sample</th><th>1-10</th><th>11-20</th><th>21-30</th><th>31-40</th><th>41-50</th><th>51-60 <th>Sample</th><th>1-10</th><th>11-20</th><th>21-30</th><th>31-40</th><th>41-50</th><th>51-60 <th>Sample</th><th>1-10</th><th>11-20</th><th>21-30</th><th>31-40</th><th>41-50</th><th>51-60 <th>Sample</th><th>1-10</th><th>11-20</th><th>21-30</th><th>31-40</th><th>41-50</th><th>51-60 </th></th></th></th></th></th>	Sample	1-10	11-20	21-30	31-40	41-50	51-60 <th>Sample</th> <th>1-10</th> <th>11-20</th> <th>21-30</th> <th>31-40</th> <th>41-50</th> <th>51-60 <th>Sample</th><th>1-10</th><th>11-20</th><th>21-30</th><th>31-40</th><th>41-50</th><th>51-60 <th>Sample</th><th>1-10</th><th>11-20</th><th>21-30</th><th>31-40</th><th>41-50</th><th>51-60 <th>Sample</th><th>1-10</th><th>11-20</th><th>21-30</th><th>31-40</th><th>41-50</th><th>51-60 <th>Sample</th><th>1-10</th><th>11-20</th><th>21-30</th><th>31-40</th><th>41-50</th><th>51-60 </th></th></th></th></th>	Sample	1-10	11-20	21-30	31-40	41-50	51-60 <th>Sample</th> <th>1-10</th> <th>11-20</th> <th>21-30</th> <th>31-40</th> <th>41-50</th> <th>51-60 <th>Sample</th><th>1-10</th><th>11-20</th><th>21-30</th><th>31-40</th><th>41-50</th><th>51-60 <th>Sample</th><th>1-10</th><th>11-20</th><th>21-30</th><th>31-40</th><th>41-50</th><th>51-60 <th>Sample</th><th>1-10</th><th>11-20</th><th>21-30</th><th>31-40</th><th>41-50</th><th>51-60 </th></th></th></th>	Sample	1-10	11-20	21-30	31-40	41-50	51-60 <th>Sample</th> <th>1-10</th> <th>11-20</th> <th>21-30</th> <th>31-40</th> <th>41-50</th> <th>51-60 <th>Sample</th><th>1-10</th><th>11-20</th><th>21-30</th><th>31-40</th><th>41-50</th><th>51-60 <th>Sample</th><th>1-10</th><th>11-20</th><th>21-30</th><th>31-40</th><th>41-50</th><th>51-60 </th></th></th>	Sample	1-10	11-20	21-30	31-40	41-50	51-60 <th>Sample</th> <th>1-10</th> <th>11-20</th> <th>21-30</th> <th>31-40</th> <th>41-50</th> <th>51-60 <th>Sample</th><th>1-10</th><th>11-20</th><th>21-30</th><th>31-40</th><th>41-50</th><th>51-60 </th></th>	Sample	1-10	11-20	21-30	31-40	41-50	51-60 <th>Sample</th> <th>1-10</th> <th>11-20</th> <th>21-30</th> <th>31-40</th> <th>41-50</th> <th>51-60 </th>	Sample	1-10	11-20	21-30	31-40	41-50	51-60							
61.0%	39.0%	100.0%	46.0%	54.0%	100.0%	41.0%	100.0%	41.0%	59.0%	100.0%	33.0%	67.0%	100.0%	62.0%	38.0%	100.0%	46.0%	54.0%	100.0%	33.0%	100.0%	46.0%	54.0%	100.0%	33.0%	67.0%	100.0%	62.0%	38.0%	100.0%	46.0%	54.0%	100.0%	33.0%	100.0%	46.0%	54.0%	100.0%	33.0%	67.0%	100.0%	62.0%	38.0%	100.0%	46.0%	54.0%	100.0%	33.0%							
Average Hit Percentage														Average Hit Percentage																																									
192.00														192.00																																									
45.25%														45.25%																																									
218.00														218.00																																									
400.00														400.00																																									
Global Average														Global Average																																									
192.00														192.00																																									
45.25%														45.25%																																									
218.00														218.00																																									
400.00														400.00																																									
extremely volatile														extremely volatile																																									

Abbildung A.1.2.: Gesamtübersicht der Tests im Spiel Rivals of Aether

A. ANALYSEDOKUMENTE

The Battle for Wesnoth

Damage										Range														
Baseline VS 70%	Baseline VS 110%		Baseline VS 130%		Baseline VS 150%		Baseline VS 75%	Baseline VS 125%		Baseline VS 150%		Baseline VS 200%												
Sample 1-10	10	9	10	7	7	10	2	8	10	9	10	10	10	10										
Sample 11-20	9	11	10	4	6	10	4	6	10	11	9	10	10	10										
Sample 21-30	8	2	10	5	5	10	3	7	10	3	7	10	10	10										
Sample 31-40	10	3	10	4	6	10	5	5	10	5	9	10	10	10										
Sample 41-50	9	1	10	6	4	10	1	9	10	2	8	10	10	10										
Sample 51-60	9	1	10	7	3	10	2	8	10	4	8	10	10	10										
Sample 61-70	7	3	10	4	6	10	4	6	10	2	8	10	10	10										
Sample 71-80	10	9	10	5	5	10	3	7	10	3	7	10	10	10										
Sample 81-90	10	9	10	6	6	10	2	8	10	2	8	10	10	10										
Sample 91-100	8	1	10	4	4	10	2	8	10	2	10	10	10	10										
Average Win Percentage	91.00%	9.00%	100.00%	48.00%	62.00%	100.00%	28.00%	72.00%	100.00%	18.00%	82.00%	100.00%	58.00%	42.00%	100.00%	5.00%	95.00%	100.00%	1.00%	99.00%	100.00%	0.00%	100.00%	100.00%
Average Win Percentage	100.00%										336.00%													

HP										Speed														
Baseline VS 70%	Baseline VS 110%		Baseline VS 130%		Baseline VS 150%		Baseline VS 70%	Baseline VS 110%		Baseline VS 130%		Baseline VS 150%												
Sample 1-10	8	2	10	4	6	10	2	8	10	1	9	10	10	10										
Sample 11-20	7	3	10	6	4	10	2	8	10	2	8	10	10	10										
Sample 21-30	9	1	10	3	7	10	3	7	10	3	7	10	10	10										
Sample 31-40	9	1	10	4	4	10	4	10	10	4	10	10	10	10										
Sample 41-50	10	3	10	5	9	10	2	9	10	5	9	10	10	10										
Sample 51-60	8	2	10	5	5	10	3	7	10	3	7	10	10	10										
Sample 61-70	9	1	10	4	6	10	1	9	10	1	9	10	10	10										
Sample 71-80	8	3	10	3	7	10	3	7	10	3	7	10	10	10										
Sample 81-90	8	2	10	4	6	10	4	6	10	3	7	10	10	10										
Sample 91-100	10	10	10	4	10	10	2	10	10	10	10	10	10	10										
Average Win Percentage	80.00%	15.00%	100.00%	41.00%	59.00%	100.00%	22.00%	78.00%	100.00%	10.00%	90.00%	100.00%	81.00%	19.00%	100.00%	43.00%	57.00%	100.00%	30.00%	70.00%	100.00%	24.00%	76.00%	100.00%
Average Win Percentage	100.00%										336.00%													

Abbildung A.1.3.: Gesamtübersicht der Tests im Spiel Battle for Wesnoth

Baseline Tests

	Totally Accurate Battle Simulator			Rivals of Aether		Battle for Wesnoth			
	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline
Sample 1-10	2	8	10	1	9	10	5	5	10
Sample 11-20	6	4	10	6	4	10	5	5	10
Sample 21-30	4	6	10	3	7	10	5	5	10
Sample 31-40	7	3	10	6	4	10	6	4	10
Sample 41-50	9	1	10	4	6	10	5	5	10
Sample 51-60	3	7	10	5	5	10	4	6	10
Sample 61-70	4	6	10	7	3	10	3	7	10
Sample 71-80	6	4	10	7	3	10	5	5	10
Sample 81-90	3	7	10	5	5	10	6	4	10
Sample 91-100	5	5	10	6	4	10	6	4	10
Average Win Percentage	49.00%	51.00%	100.00%	50.00%	50.00%	100.00%	50.00%	50.00%	100.00%

Abbildung A.1.4.: Gesamtübersicht der Tests von Grundeinheiten gegeneinander aller Spiele

# Selbstständigkeitserklärung

Hiermit erkläre ich, daß ich die vorliegende Arbeit selbstständig angefertigt, nicht anderweitig zu Prüfungszwecken vorgelegt und keine anderen als die angegebenen Hilfsmittel verwendet habe. Sämtliche wissentlich verwendete Textausschnitte, Zitate oder Inhalte anderer Verfasser wurden ausdrücklich als solche gekennzeichnet.

Mittweida, den 20. September 2023

A solid black rectangular box used to redact the signature of the author.

---

Tim Riemann