Mathematical balance metrics in competitive multiplayer games

Godde Godde

Department of Computer and Systems Sciences

Degree project 15 HE credits
Degree subject: Computer and Systems Sciences
Degree project at the Bachelor level
VT 2014 2014
Advisor: P
Reviewer:
Swedish title:

Stockholm University

Mathematical balance metrics in competitive multiplayer games

Godde Godde

Abstract

This thesis introduces simple mathematical balance metrics that can be used to evaluate a games' balance. Concepts such as Metagame, Balance and "Starting conditions" are defined from published papers and by documents released by prominent game developers such as David Sirlin. By generalising basic attributes for unit stats with measurements such as time, ingame cost and damage output a general balancing approach is formalized. The metrics assumes that the simulation or unit behaviours are in a whitebox where the stats and metrics used actually apply to real ingame simulation or that the behavior of the real ingame simulation have been understood enough to generalize the effects of the simulation. Several terms of balance design and balance metrics are defined.

Keywords

Game design, balance, RTS, Real-Time Strategy

Contents

1. Introduction

The subject of study is of interest to the players who want to understand and use the game balance to their advantage against live opponents or AI opponents. As competitive gaming becomes more common, the need for game balance evaluation grows. It is also of interest to game designers who strive to deliver a game where balance is important.

Balance design is of great importance to the longevity and depth of competitive games(Alexander and Perlman, 2014; Jaffe, 2013; Sirlin, 2008a). Balance can include concepts such as fairness, viability of options and depth. However clear guidelines for how to create a balanced game is lacking in the scientific literature and building up a method of balancing games by using mathematical analysis on basic unit characteristics is also lacking. Game balance design can be approached from many different angles: Statistics from real games. User feedback. Mathematical analysis of ingame parameters. By defining and setting up guidelines of how to perform basic unit balance comparisons, an elementary understanding of game balance can be made which can then be built upon by future research.

1.1 Motivation

Game mechanics are the rules of the game. Game mechanics in computer games are inherently mathematical. All computer games can be analyzed mathematically but many games are very complex. The game mechanics in Planetary Annihilation, like many Real-Time-Strategy games, requires a lot of computation. By defining simple mathematical metrics for the game we can derive models that requires much less computation and are possible to analyse from a game design balance perspective.

1.2 Research question

What kinds of basic mathematical metrics are there for units such that those metrics can be derived by the game mechanics?

2. Extended Background

Because of the complexity of Real Time Strategy games, which is further amplified by the simulation aspect of Planetary Annihilation, strategies are constantly found and developed. There are some strategical choices that can be considered true simultaneous moves or blind choices. As the game starts the players don't know where the enemy opponent have started and they also don't know which initial factory the opponent have chosen which have some strategical ramifications. If these strategic ramifications can have critical influence on the outcome it could be considered blind Rock-Paper-Scissors as these simultaneous moves can decide the outcome of the game before the opponents have even had a chance to scout each other. This have can various consequences for the metagame. These blind strategic choices are likely to have an effect on the game but it is unclear as of this writing how strong influence this have on the outcome of the individual game and to the metagame between 2 players that know each other and 2 players that don't know each other. It could be likened to a weighted Rock-Paper-Scissors where the initial choice have a likeliness to favor 1 side like it could also be likened to the metagame of a fighting game as described by Jaffe (2013) or a strategy game with multiple sides where the initial choice of character or race increases the probability of 1 player winning or losing. If such a relationship exists then there should also exist a Nash-equilibrium(Nash et al., 1950) of strategies that min-max the outcome for the individual player.

Metagame, starting conditions, pre-game choices and pre-game choices

Metagame for Starcraft and hereby adapted to Planetary Annihilation could be described as "a complex interplay between the game community and the game itself" (Carter et al., 2012). "Metagame" as described by the popular online Starcraft wiki, Liquipedia:

1. Preparation done before a match to exploit current trends in Starcraft.

2. Preparation done specifically to exploit an opponent's or map's style of play.

3. Strategic decisions designed specifically to exploit a player's reaction or weakened mental state in the future. These are also known as 'mind games' or 'psychological warfare'.

Crawford (1984) argues that symmetric games have several inherent weakness where simplicity is the greatest one. Crawford argues that a dominant strategy or a strategy that promises to be truly effective, as he puts it, will be used by both sides and this makes the game be more about execution than strategy. However even in asymmetric games there might be dominant strategies for some or all match-ups. Since Planetary Annihilation only have 1 match-up it is interesting to see how the game is balanced to avoid 1 dominant strategy. Crawford further argues that a nontransitive or triangular relationships is the strength of asymmetric games like the Rock-Paper-Scissors game. However Alexander and Perlman (2014); Jaffe (2013); Sirlin (2012) argues that just such a relationship is bad for strategy game as the outcome is determined even before the game starts. It might be part of the metagame and it most likely will be a part of the metagame if there are asymmetric choices pre-game that locks the player into a certain race or character(Jaffe, 2013). However I will extend the definition of blind RPS further than Crawford and argue that such relationship is bad for the game as the players have no ability to scout each other and the game can be decided by a blind RPS decision. Triangular relationships once the game has started are desirable however. removing the distinction between (Sirlin, 2012).

The metagame is specific to the community from which the players themselves derive from. Isolated communities or player bases are likely to develop their own specific strategies which constitute the metagame. In a way the metagame between top competitive players incorporate the pinnacle of play(Jaffe, 2013). Once a metagame has been established it can then be analysed as Jaffe (2013) proposed by finding the Nash-equilibrium for character selection or the choice of strategy when he analysed what he called "Metagame balance". Theoretically the strategy match-ups in an RTS could be likened to the character select in a fighting game but the strategy match-ups in an RTS are so fluid and vast that the number of match-ups is enormous unless the match-ups are clearly defined.

3. Method

I have analysed a game under development, Planetary Annihilation and I have developed and applied mathematical balance metrics in pursuit of finding and evaluating the viability of the different options that are available to the player. A case of Design Science to develop and evaluate an artefact.

Design science

Design research creates novel artefacts as well as knowledge about them, their use, and their environment. Design science is a special strand of design research that aims to create artefacts that address problems experienced by people(Johannesson and Perjons, 2012). Planetary annihilation is a stochastic zero-sum strategy game played in real time. That is to say that the players makes simultaneous moves without knowing the exact strategy of the enemy opponent. It is currently under development.

The strategies deployed by the player have to be considered stochastic strategies as Planetary Annihilation is a game of imperfect information where the opponents' strategy rarely are fully known. Game balance design can be approached from many different angles. Playtesting. Statistics from real games. User feedback. Mathematical analysis of ingame parameters.

Arguably playtesting is the best method of determining game balance(citation needed). However playtesting takes a lot of time and requires that people of required skill can participate.

Mathematical unit balance metrics can be applied without the need for playtesting and can easily be automated by reading unit stats directly from the gamefiles.

The goal of a metric is that the relationships defined by the metric should be the same regardless of the unit of measurement. If we use a metric such as Damage Per Second and compare relationships between units, the relationships should not change if we instead measure it as Damage Per Minute.

A unit relationships/ratio is defined as Unit A metric stat divided by UnitB metric stat. Once divided, the ratio should be the same regardless if we measure damage over time with Damage Per Second or Damage Per Minute.

In a way those relationships that are found and described is highly dependent on skill of the researcher implementing those metrics. Defining the balance of the game and understanding the balance of the game gives the players and designers more knowledge in a descriptive way which can be spread more easily. The progress on mathematical balance metrics from 1 person or 1 community can transcend into the player-base and further the balance process of the game as new strategies and tactics are discovered. Personally I have seen and experienced that mathematical ventures into defining the balance of a Planetary Annihilation have been meet by critique on the game forum that the game is simply to complex to be analysed mathematically. Sirlin (2008b) agrees that mathematical analysis is too complex and too time consuming to apply for complex games even for

fighting games which could be considered much less complex than RTS. Jaffe (2013) argues that Monte Carlo Tree Search drastically reduce the complexity and might be viable as finding a rough balance by mathematical analysis without needing humans to perform the actual playtesting. Players intuitively develop models of their own describing the balance of the game. Those models that can be easily described and defined by players can most likely be subject to mathematical analysis which also makes them more accessible to other players and game developers. Defining such relationships have to be done on a game to game basis but common relationships can be applied between different games.
I have used a script written in Python to create HTML pages that I have uploaded to a server. Once on the website, the values of the metrics are basically available to the whole playerbase of Planetary Annihilation. I have used numbers that used by the game engine. Since I am only using the mathematics in idealized conditions so the actual correlation to the ingame simulation is unclear. As the game is under development, the simulation could rapidly change even though the script can search the game files for changed values in the unit files. There could also be bugs in game software that makes the actual output unclear.
RTS is particularly difficult to even approach an approximate solution as the statespace and actionspace are superexponential. Other methods could apply evolutionary AI searching techniques which tries different inputs and choses ones that have better output. However the actionspace is enormous in RTS and the interaction between divergent strategies that is required to adhere to the metagame is hard to pinpoint. Even if the AI evolves to use the dominant strategy, it might still not be able to coevolve divergent strategies which would be required if the AI is to develop a set of mixed strategies to satisfy a Nash-equilibrium. If the dominant strategy as a pure strategy doesn't satisfy a Nash-equilibrium that is. However if a pure strategy is found to satisfy a Nash-equilibrium then the developers are likely to change the balance of the game in order to remove this dominant strategy which have no counters as this clearly violates the viability of different options and solves the game.

4. Results and Analysis

"The designer had already built a series of small playable prototypes, each aimed at one sub-part of the game: the economic system, the combat system, and base building and base defense scenarios."(Nelson and Mateas, 2009)

Supreme Commander, Total Annihilation, Zero-K uses the same streaming model as Planetary Annihilation. Planetary Annihilation uses a a so called "streaming" resource model. This makes the income easier to determine as comes at a constant flow depending on the resource generating units compared to say, Starcraft where resources are delivered at discrete points in time as units deliver resources back to the main buildings!?.
However, the economic model in Starcraft can be modeled as a stream of income as was done by Churchill and Buro (2011)

In close-combat.
Unit match-up.
Combo match-up.
Density case.
Time to live.

4.1 Basic unit stats

What is referred to as "unit stat" or "unit statistics" in computer games are in many cases unit properties rather than statistically measured values. These "unit stats" are usually derived directly from the game files or ingame tooltips and in this thesis I will refer to them as "unit stats".

Basic unit stats in PA:

Weapon
I define weapon as a device on a unit for which the whole purpose is to damage or destroy enemy units. Most units in the game have 1 or several weapons. These weapons might operate in unison or individually from each other where the different weapons might have different ranges and/or target different types of enemy units only such as anti-air weaponry or torpedoes against ships.

Damage
Damage is the usually an integer value which constitutes how much damage each projectile or each laser beam does.

Weapon Reload
This is how fast the weapon reloads measured in seconds.

Damage Per second
Damage Per second, DPS, is a widely used metric used by gamers. It is measure of how much damage the weapon does on average every second. The calculation is as follows:
Damage / Reload
To get the total DPS output for the unit in question, the DPS of the different weapons on the unit can simply be added together.

Health
Health is also called Health Points. All units have Health which is usually represented by an integer value.
Damage usually trade with a 1:1 ratio for Health with few exceptions.

These terms can be used for a many games in the RTS genre in games such as Starcraft, Supreme Commander, Zero-K and Total Annihilation. However there have to be considerations taken of damage modifiers which are usually in the form of types armor, weapon types or status effects on weapons and units. Those damage modifiers have to be accounted for on a case by case basis as unit match-ups for example.

Suppose we have 2 units with continuous damage output. Modeling a battle between these units can be done as such:

StrengthOfX = DPSofX * HealthOfX

StrengthOfY = DPSofY * HealthOfY

StrengthOfX - StrengthOfY where negative values means victory for unit Y and positive values means victory for unit X. The impact of the DPS of the unit is directly proportional to the Health of the unit.

4.2 Basic resources

Metal
Metal is a global resource and is streamed continually from the Commander, Metal Extractors and Advanced Metal Extractors and can also be reclaimed from the wreckage of destroyed units. The amount of metal that can be accumulated and stored is limited but can be increased by making Metal Storages.

Energy
Energy is a global resource and is streamed continually from the Commander, Energy Plants and Advanced Energy Plants and a few other units. The amount of energy that can be accumulated and stored is limited but can be increased by making Energy Storages.

Build Power
Build power is a local resources as it have to be provided by factories and fabrication units. Build Power is the potential metal output per second which is provided by factories and fabrication units. Initially the player is limited to their Commander but will soon build factories which can in turn produce fabrication units.

Cost
Cost can be normalized to metal cost. All units in the game have a metal cost. How quickly a unit can be built depends on the Build Power that have been assigned to the project as their combined Metal output per second determines how fast the unit is finished.

Build energy drain

Build energy drain is how much energy each factory or fabrication unit needs to use their Build Power. If the player have less energy than required to sustain all energy needs then factories and fabrication units will also have a reduced Build Power.

Infrastructure cost

Infrastructure cost is how much metal that is required to make a factory or fabrication unit and provide energy to keep it fabricating. Infrastructure cost is calculated in metal. Infrastructure cost is variable on how cheap and cost-effective energy plants are. If energy production is cheap then the infrastructure cost of running a factory or fabrication unit is also lower.

Strength per 1 metal

Strength per 1 metal is a general unit stat that informs about the general combat effectiveness for cost.

(DPS / cost) * (Health / cost)

Overkill factor

Overkill is important to consider in unit match-ups. The lower bound of overkill is how big percent of the damage that will be wasted even if all projectiles hit the target. Another pretty accurate lower bound is how much percent of the damage that will be wasted if several weapons whose targeting can't be split among several targets always fire simultaneously. While some of the projectiles might hit other units after their intended target is destroyed, depending on the distance and spread between the enemy units and the weapon characteristics this might be unlikely and the rest of the salvo damage is wasted.
The higher bound of overkill can potentially be unlimited as many units can fire at a single targets where only few of the projectiles will hit before the target is destroyed and the rest of the damage is wasted.

Area of effect

The effect of weapons with Area of Effect also called "AoE" or "splash" depends on the density of enemy units. The DPS can potentially be multiplied many times if the Area of Effect from the weapon can hit several units at once.

Roll-off time

Factory roll-off time is important in order to determine the smallest amount infrastructure requirements for specific units.

Cumulative strength

The cumulative strength of a several units

SUM = UnitHealth*DPS*1 + UnitHealth*DPS*2 + UnitHealth*DPS*3 + ... + UnitHealth*DPS*n

dA/dt = -Kd*D dD / dt = -KB*A

$$\sum_{i=1}^{n} i = Health * DPS * n$$

Groups versus groups

By subtracting the cumulative strength from each other. With overkill the formula can be described as SUM = As numbers grow the formula will more and more adhear to Lanchasters square law(**?**) and will look like this:

Overkill

Overkill is the amount of damage that is lost due to that the projectile damage exceeds the amount of the targets Health. Overkill may apply both when the projectile damage exceeds the maxium health of an intended target and when the target have less health left when the projectile hits the target. Overkill is also important to consider when the weapons fire several projectile in a volley. Even if the overkill from individual shells might be minimal, the overkill due to that the target is already dead when the projectiles reaches the target, might be large because the rest of the volley might not do any actual damage to any other unit. If several units fire in unison the effects of volley overkill can still apply even if the weapon only fires 1 projectile at a time. Overkill is important to consider when units with very high damage projectiles face off against units with very low health.

Bottleneck resource

There are 3 main resources in PA. Metal, energy and buildPower. The availability of these resources varies at different points in the game and depending on the map. Maximising economic growth or other production requires a balanced income of all required resources. At any point in time 1 resource can be considered a bottleneck resource. Typically Metal is the bottleneck resource in PA as metal production is much more limited than energy or buildPower production.

Economy Time Cost

EconomyTimeCost is a metric to derive the relative time required from the economy in order to produce a unit. In its' simplest form it is % of bottleneck resource production * Time. However storing resources makes it more complicated.

Interception angle

The angle which a unit can intercept an incoming unit.

Interception area

The area that unit can protect by intercepting a single unit.

Starcraft random overkill prevention

Damage is checked whether a unit has received enough damage to be killed. The order in which units are tried might be tried and units far back might not reach an enemy at the back of a row of units when all other targets are considered dead.

Image of strength per 1 metal over metal cost

5. Discussion

Further research can go into defining the possibilities of: The difficulty of breaking down input into approximate discrete strategies is big but it is something that players and developers can still do so that the metagame of a such complex game as PA can be analysed to find potential mixed strategies that satisfy a Nash-equilibrium. By analysing the game mathematically and finding discrete strategies bg While the set of discrete strategies might not contain The set of discrete strategies Further research can define at unit effectiveness with micromanagement as making player input and attention as a resource.

6. Glossary of terms

In order of appearance?
Balance
In this thesis, balance is defined as fairness, viability of options and depth in combination.
Depth.
A multiplayer game is deep if it is still strategically interesting to play after expert players have studied and practiced it for years, decades, or centuries.(Sirlin, 2008a)
Blind RPS(Rock-Paper-Scissors)
Blind RPS involves blind guessing, aka a stochastic approach.
Stochastic
Stochastic Strategy
A stochastic strategy means choosing a strategy with limited information and can be likened to simultaneous moves in game theory.
When the choice of strategy is based on the likeliness of your opponent choosing a specific strategy.
Zero-Sum game

Unit stats

References

Alexander, S. and Perlman, M. (2014). Competitive gaming.

Carter, M., Gibbs, M., and Harrop, M. (2012). Metagames, paragames and orthogames: A new vocabulary. In *Proceedings of the International Conference on the Foundations of Digital Games*, pages 11–17. ACM.

Churchill, D. and Buro, M. (2011). Build order optimization in starcraft. In *AIIDE*.

Crawford, C. (1984). The art of computer game design.

Jaffe, A. B. (2013). *Understanding Game Balance with Quantitative Methods*. PhD thesis.

Johannesson, P. and Perjons, E. (2012). A design science primer. *Unpublished Manuscript, February*, 25.

Nash, J. F. et al. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, 36(1):48–49.

Nelson, M. J. and Mateas, M. (2009). A requirements analysis for videogame design support tools. In *FDG '09: Proceedings of the 4th International Conference on Foundations of Digital Games*, page 137–144, New York, NY, USA. ACM, ACM.

Sirlin, D. (2008a). Balancing multiplayer games, part 1: Definitions. *Sirlin. net*, 17.

Sirlin, D. (2008b). Balancing multiplayer games, part 4: Intuition. *Sirlin. net*, 17.

Sirlin, D. (2012). A discussion of balance.

Stockholm University
Department of Computer and Systems Sciences
Forum 100
SE-164 40 Kista
Phone: +46 8 16 20 00
www.dsv.su.se