

# Levelling the Playing Field – Games Handicapping

CLIVE CHANDLER & LEN NORIEGA

Faculty of Computing, Engineering and Technology

Staffordshire University

P.O. Box 334, Beaconside, Stafford ST16 9DG

UK

<http://www.fcet.staffs.ac.uk>

*Abstract:* - Imagine the look on your ten year old sons face having just been totally annihilated by an over zealous big brother / father combination in Halo-2. Ignoring the fact that he may be scarred for life by this experience, his whole attitude towards the enjoyment of multiplayer games is tarnished.

This is the experience of many multiplayer gamers, they purchase a game and having gotten to grips with the key-strokes and combination moves decide to join an online game server. They are completely destroyed by players who have been playing the game for the past three week's non-stop, no eating, no drinking just playing the game, the so-called *game experts*. This puts the whole regime of online community in jeopardy.

The purpose of this paper is to analyse possible algorithms to develop a handicapping system which would allow a *newbie* to compete with an *expert* in a multiplayer environment and both still achieve a satisfying game. In addition a series of simulated handicapped Halo-2 games are analysed to assess one form of handicapping and provide a forum for discussion of other approaches with experienced players.

Finally, a new algorithm is postulated which would provide suitable co-operative play without losing challenge.

*Key-Words:* - Handicapping, Games, Turing, Algorithms, Fair Play, Handicap, Passive Imposed, self regulated

## 1 Introduction

Imagine the look on your ten year old sons face having just been totally annihilated by an over zealous big brother / father combination in Halo-2. Ignoring the fact that he may be scarred for life by this experience, his whole attitude towards the enjoyment of multiplayer games is tarnished.

This is the experience of many multiplayer gamers, they purchase a game and having gotten to grips with the key-strokes and combination moves decide to join an online game server. They are completely destroyed by players who have been playing the game for the past three week's non-stop, no eating, no drinking just playing the game, the so-called *game experts*. This puts the whole regime of online community in jeopardy.

Some gaming communities have attempted to compensate for this over abundance of experts by creating novice or intermediate only servers. The

problem is how do you identify a novice or an intermediate?

These sites become a breeding ground for the experts to slaughter unwitting gamers who believe this version of the game world is populated only by novices. The expert becomes the white shark of this game, preying on the unwitting and "*levelling up*" their character at the expense of the "*newbie*".

Handicapping is one route forward, it has so far not gained popularity with the gaming community who argue that any restriction on their gameplay is a restriction on their freedom of expression!

## 2 Origins of Handicapping

The origins of the word *handicapped* go back to the eighteenth century, and apply to the handicapping of horses by adding weights to even out a contest.

As the *Oxford English Dictionary* puts it, it would then have meant “*any encumbrance or disability that weighs effort and makes success more difficult*”

Amundson <sup>(1)</sup>, argues that the basis of handicapping comes from a dramatic story about war veterans and a king who only allowed people with impairments to beg in the street, known as “*cap-in-hand*”.

In terms of games the main goal appears to be make the game more realistic in terms of human opponents e.g. games developed since the 1950’s to beat a human at chess <sup>(2,3,4)</sup>, or the more complex game of *Go* <sup>(5,6)</sup>

As such most considerations of game levelling or handicapping have been involved with the implementation of suitable golf handicapping algorithms <sup>(7,8)</sup> rather than player or game play equality.

There appears to be a dearth in academic literature in the specific area of levelling the gap between expert and novice game players (handicapping). A number of researchers have investigated the development of game playing communities on-line or the behaviour of peer groups and on-line players <sup>(9, 10)</sup>. However handicapping has somewhat been ignored.

### 3 Game Turing Test

In traditional IT the Turing test proceeds as follows; a human judge engages in a natural language conversation with two other parties, one a human and the other a machine; if the judge cannot reliably tell which is which, then the machine is said to pass the test. It is assumed that both the human and the machine try to appear human. In order to keep the test setting simple and universal (to explicitly test the linguistic capability of the machine instead of its ability to render words into audio), the conversation is usually limited to a text only channel such as teletype machine as Turing suggested. As yet no system has managed to succeed in this endeavor. This test was first developed by Alan Turing in the 1950’s <sup>(11)</sup>

In games terms if a player can play in some multiplayer games arena and they cannot tell after a suitable elapsed time whether they are competing with a “*Bot*” or a human player then the games AI will have successfully fulfilled the Turing criteria. This we can term the *Game Turing Test (GTT)*.

### 3.1 Turing Handicapping

Given the discussed criteria for a Game Turing test, a suitable expansion would be to adapt this test to illustrate successful criteria for any form of handicapping adopted. Thus if our player cannot ascertain or determine whether their opponent, in either a death match scenario or a multiplayer arena, is a novice, intermediate, or expert then it can be argued that the system has passed the *Turing Handicapping Test (THT)*.

### 4 Current Handicapping Games

As mentioned in section 3 the majority of handicapping relies on the more traditional games such as Golf, Horseracing and other sports games provide a direct translation from real world handicapping systems to the computer generated sports game. Anecdotal forum discussions carried out by the author with groups of game playing students tend to support this style of handicapping as expected and accepted. The argument from players is that the goal of the game is to progress or “*level up*” by improving the players’ golf handicap or winning a race despite the handicapping imposed on the players’ horse.

However the forum group did not show approval or support for any form of handicapping in Role Playing games (RPG) or First Person Shooters (FPS), although it was unclear as to why.

Due to the perceived unpopularity of modern games players to support some form of handicapping there are few modern FPS or RPG games which have implemented any form of handicapping at all. One of the more popular X-Box games *Halo-2* (see Figure 1 below) has implemented a self-imposed and selected handicapping. Here the player selects either a Low, Medium, or High handicap for their own player (low, moderate, severe).



Halo-2 Death match

Even this simple self-imposed handicapping is not well received, and in many cases it is an undocumented feature.

Bungie developers *Max Hoberman* (Design Lead responsible for multiplayer, UI and Xbox Live design for Halo-2) and *Jo Tung* (Designer Multiplayer Halo-2) in a recent discussion on the Bungie Halo-2 news web-site<sup>(12)</sup> stated the following:

*“Jo Tung cruelly notes that he NEVER uses handicapping, and insists on playing his seven year old niece to make himself feel more ‘awesome’ “*

*“Max Hoberman says, “I’ve used the handicap feature when I’m playing with a small group (4 or less) with mismatched skills. Usually split screen, but sometimes networked. I think I’ve always been the best player in these scenarios, so I set a handicap on myself to even the playing field. The setting I use depends on the skill and number of my opponents, but it’s usually moderate or even severe. Severe is damn hard though, especially playing Rockets! Anyway, I usually end up playing a Slayer variant, some FFA and sometimes me vs. everyone else. The other players have fun because they’re getting kills and not just getting reamed, I have fun because it’s more challenging to stay alive and to get kills. I’ve even played where multiple people had handicaps, roughly equivalent to their skill levels, and the new player ended up winning the game as a result”.*

Although these comments are from a Bungie news web site it does show the indicative nature behind the idea of handicapping, the author of the site argues that this feature is hopelessly underused.

Currently this is the only First Person Shooter (FPS) offering which offers this kind of handicapping approach. It is designed to reduce the shield strength of the player by 25%, 50% and 75%. As such this self-imposed handicapping does not appear to be very successful.

More recently there has been an offering by the Disney Studio organisation “*Kingdom Hearts II*” (see Figure 1) argues to impose a new style of handicapping to level the playing field, as yet there are no available details.



Figure 2: Kingdom Hearts II

## 5 Simulation Tests

In order to get further insight into handicapping approaches and player opinion a series of simulated multiplayer death match scenarios were conducted utilising *Halo-2* with a varying degree of handicapping imposed. Figure 3 illustrates the lab set-up.



Figure 3 X-Box multiplayers Simulation

Each of the players (5 in Total) was a student studying a Masters degree in Games Design and expert Halo-2 game players. In addition the author was added as a novice player to ascertain the efficacy of the handicapping system.

Six series of handicapping tests were employed i.e.

**Test 1: No handicapping 40 minute series of death match – ranking order established**

**Test 2: Low handicapping (25% Shield loss) imposed on each of the players in ranking order**

**Test 3: Moderate handicapping (50% Shield Loss) imposed on each of the players in ranking order**

**Test 4: Severe handicapping (75% Shield Loss) imposed on each of the players in ranking order**

**Test 5: All players were moderately handicapped except the novice player**

**Test 6: Return to no handicapping to ascertain effect of handicapped play**

In addition a forum discussion was carried out to ascertain if the THT (Turing Handicap Test) criteria were perceived to be a success and to gain player insight and opinion.

### 5.1 Simulation Results

In most of the six series of tests employed the ranking order appeared to be maintained with the handicapped player achieving the lowest score. One notable exception was that of the so-called novice player, in some of the scenarios, approx 25% did not rank above the handicapped player. Indeed even when all players were handicapped in Test 5 the novice player was ranked 2 out of 6. One argument for this was that the handicapping system or imposition could be circumnavigated by developing different game strategies and as such the handicapped player “learns” to overcome the handicap.

Interestingly one player developed a different learning style once handicapped and employed this “new” strategy in the final series where no handicapping was employed, achieving a higher ranked system.

It was noted by all the players that the system itself was “too severe” even at its lowest setting and that far from improving the game for the player thus handicapped, it in fact caused them to lose interest and find alternate amusement.

One interesting result was the self-imposed weapon ban. During play it was found that whoever utilised the rocket launcher gained a high degree of dominance as did the player who utilised the flyers present in some scenarios. In order to remove this dominance the group decided that the death matches would be played with a ban on flyers and the rocket launcher.

## 6 Handicapping Possibilities

Given the forgoing scenarios and subsequent discussions a number of possible handicapping

approaches were considered, refined and presented to the forum group for their opinion. The algorithms fall into two categories, passive imposed (PI) and self Regulated (SR).

### 6.1 Passive Imposed (PI) Algorithms

These algorithms are discussed in principle, and apply to a direct imposition of players skill sets based upon some form of measured metric such as ranking, accuracy etc.

#### 6.1.1 Shield Strength

One possible algorithm discussed was that of shield strength; Halo-2 utilised a self-imposed reduction in shield strength which, as indicated by scenario testing, proved both unpopular and relatively too severe for success. Similarly this approach offered no real possibility for THT success.

Further modification of this algorithm would be to utilise an imposed increase in the shield strength of the Novice player; identified after a series of games i.e. the lowest ranking player would have the highest shield strength. Players felt that this would again be too severe.

An alternate algorithm would be by utilising a scaled shield strength approach i.e. similar to the handicapping in horse racing an algorithm based on number of kills and experience (rather than number of wins and weight) should determine the shield strength of the player. This would seem to have some support and the possibility of THT success.

#### 6.1.2 Accuracy

In order to handicap the more expert players an algorithm to impose a hit restriction was devised i.e.

Most modern games have an initial tutorial, for example *Medal of Honour Allied Assault*, see Figure 4, whereby the new player is taken through the controls and is shown how to fire a number of the weapons available in the game. These scenarios usually consist of firing at targets and hitting a number of bull’s-eyes or destroying a decoy.



Medal of Honour Allied Assault

An algorithm can be utilised whereby the accuracy of the player is calculated during this weapons training mission and by suitable handicapping the most accurate player hits the opponents the least number of times.

Presenting this approach to the forum yielded mixed reactions; firstly 70% thought that this approach might work but that it could be fooled i.e. once a player found the “secret” of the algorithm they would only need to “miss” a number of targets to gain a higher accuracy rating. Of the dissenting group most felt that it would lead to a loss of challenge and interest for the handicapped players and felt that any form of handicapping which restricted the experts strategies would create bad gameplay.

### 6.1.3 Hit Points

One final approach was considered i.e. that of manipulating the hit points required in order to kill an opponent. Again the idea could be based upon accuracy or ranking of the expert player. Here however rather than affecting the shield strength a modification of the weapons carried was considered; If the most skillful player fires a standard weapon at a novice opponent it would take X accurate shots to kill the opponent; However, if the novice player fires the same weapon at the expert player it would take X-y accurate shots. In terms of Dungeons and Dragons strategy the novice player would have a larger number of hit-points than the expert player.

This again had some support but it was felt that game players would either find a cheat code; this would seem to be the drive in online game playing for example <sup>(13)</sup>, or fool the algorithm during measurement.

### 6.1.4 PI Overview

On the whole Passive Imposed algorithms, such as those presented present a serious flaw in their implementation. Given that the nature of these game players is to gain advantage by circumventing the normal game rules if possible, any algorithmic approach to handicapping which seeks to impose a restriction of some kind on the expert player, or even to promote the novice player, pushes them to find a way around the restriction.

In other words, unless the players “buy-in” to the handicapping or restricted play, they will cheat and the level playing field would not be achieved.

Similarly although it might be possible to develop a suitably complex algorithm, it is doubtful that the players would be fooled enough to pass the THT criteria for success.

## 6.2 Self Regulated Algorithms

The self regulated approach would at the outset appear to be the more acceptable. Players either impose a handicap on themselves, such as in Halo-2, or they are fully aware of the handicapping and adhere to it, as in golf or horseracing approaches.

### 6.2.1 Self Imposed

Halo-2 as mentioned in previous sections adopts a self-imposed shield strength reduction. Also as mentioned in scenario testing this approach does not appear to have much success amongst players, mainly due to its severity. It could be argued that the severity of the shield strength could be adapted to reduce the impact, however as yet this is untested.

### 6.2.2 Ranked Handicap

Most players accept the ranking and handicapping imposed in golf games. In the opinion of the author this is partly due to the head-to-head match play approach of the genre i.e. the goal is to get a ball in a hole in as few strokes as possible, although you are attempting to beat the opponent it is your skill at getting the ball in the hole which is primary. In death match FPS games for example it is you or your opponent head-to-head but if he misses you and you don't miss him the game is over! Here the primary goal is to kill your opponent even if you trick him, sneak up on him etc. Therefore the ranking is not as direct.

However it is possible to utilize ranked tables to form a handicap, similar to team tables in games such as *Battlefield 1942* or *World of Warcraft*. Whilst successful however this does not allow the novice to play against the expert with any degree of equality and this approach seems doomed to fail.

### 6.2.3 Balanced Co-operation

It is the conjecture of the author and the premise of this paper that if somehow we can create a situation of co-operation between an expert and a novice that the novice can improve and both expert and novice can achieve an even playing field and an enjoyable game. To this end, a Balanced Co-operation algorithm is proposed; In order to allow both expert and novice to play

on a level playing field some form of metric should be utilised to determine ranking. It is then proposed that during game play the leading player should be restricted by for example number of points scored, levels gained, number of kills etc. Such that they can only be X above the lowest player.

This then would force our expert player to consider the ongoing score or position such that it would promote a co-operation between the top player and the lowest player.

## 7 Conclusions

1. Handicapping in Modern Games is still in its infancy
2. A proposed metric utilizing a modified Turing Test is proposed
3. Halo-2 utilises shield reduction which in simulation testing proved too severe to be successful
4. Suggested handicapping algorithms are either Passive Imposed or Self Regulated
5. Passive Imposed Algorithms are the least popular of those proposed
6. Ranked shield strengths held some merit and the possibility of THT success
7. Proposed Balanced Co-operation held the most merit amongst expert players

### References:

- [1] Amundson R., about the meaning of Handicap, Dept. of Philosophy, University of Hawaii at Hilo, 2006 (available at <http://www.uhh.hawaii.edu/~ronald/HandicapDefinition.htm> )
- [2] Samuel A., Some studies in machine learning using the game of checkers. IBM Journal of Research and Development. 3(3):211-229, 1959
- [3] Schaeffer, J. et al, A World championship calibre checkers program, Artificial Intelligence. Vol 53, No 2-3, pp 273-290, 1992
- [4] Schaeffer, J., The Games Computers (and People) Play, Academic Press, Vol 50 (ed Zelkowitz M.V.), 189-266, 2001
- [5] Chellapilla, K., Fogel, D., Evolving neural networks to play checkers without expert knowledge, IEEE Trans on Neural Networks, 10(6): 1382-1391, 1999
- [6] Fogel, D. Blondie24: playing at the edge of AI, Morgan Kaufmann Publishers, 2001
- [7] Scheid, F., The search for the perfect handicap, Proc. 10<sup>th</sup> Conf. on Winter Sim, Vol2, 1978
- [8] Scheid, F., Golf Competition between individuals, Proc.11<sup>th</sup> Conf. on Winter Sim, V2, 179
- [9] Manniman, T., Virtual Team Interactions in Networked Multimedia Games, Proc. Presence 2001, Phil USA, 2001
- [10] Chandler, C, Peer Group Behaviours: Social Narrative of Game Play”, WSEAS Trans on Comp, V4:10, 2005
- [11] Turing, A.M., Computing machinery and intelligence, MIND, Vol 59, No 236, Oct 1950
- [12] Bungie News Site Visited July 2006 <http://www.bungie.net/News/TopStory.aspx?story=weeklydec06>
- [13] Yan, J., A Systematic Classification of Cheating in Online Games, Netgames’05, Oct 2005