Is it Percival time yet?: A preliminary analysis of Avalon gameplay and strategy

Yuzuko Nakamura
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213
yuzi@cs.cmu.edu

ABSTRACT

The Resistance: Avalon is a hidden-roles-style board game. In this paper, we use data collected over dozens of Avalon games to make recommendations on role sets and game sizes that maximize the game-playing experience. We also evaluate the effect of various strategies on good and evil's chances of winning.

KEYWORDS

Board games, hidden-role games, game design

ACM Reference format:
Yuzuko Nakamura. 2017. Is it Percival time yet?: A preliminary analysis of Avalon gameplay and strategy. In *Proceedings of SIGBOVIK, Pittsburgh, PA USA, April 2017 (SIGBOVIK'17)*, 6 pages.
DOI: 10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

The Resistance: Avalon [1], like Mafia, is a multiplayer game centered around hidden roles. Hidden role games involve players being randomly assigned roles that are not revealed to other players. These games often feature two or more sides with their own win conditions; in particular, there is frequently an evil or sabotaging side that attempts to bluff and win people's trust in order to win the game, and a good (but generally information-less) side that must correctly guess who to trust in order to win the game.

Avalon is a two-team game themed around King Arthur: the loyal knights of King Arthur (good team) attempt to succeed three quests (missions), and the minions of Mordred (evil team) attempt to be placed on the missions so as to sabotage them and lead three of them to fail. As such, the aim of evil is to be trusted by good players and the aim of good is to determine who can be safely trusted to be sent on a mission. In addition, all members of the evil team know each other (with possible exceptions).

Avalon in its simplest form[1] features two special roles, Merlin and the Assassin. The addition of two other special roles, Percival and Morgana, creates more opportunities for bluffing, trust, and strategy so we are interested in games with these four roles:

- Merlin (good): The player with this role knows all evil roles. They must guide other good players to this knowledge. However, Merlin must be subtle in the way they do this due to the role of the Assassin.
- Assassin (evil): If good successfully passes three missions, the player with this role gets to choose one good person to assassinate (after discussing with the rest of the evil team). If the assassinated player was Merlin, evil snatches victory from the jaws of defeat and wins the game. As such, this person serves as a check on Merlin's ability to help the good team.
- Percival (good): The player with this role knows who Merlin is. As such, they can also gain indirect information about who to trust by quietly observing Merlin's actions and can appear to evil as a decoy Merlin. However, due to the role of Morgana, Percival must first determine which Merlin to trust.
- Morgana (evil): The player with this role appears as a second Merlin to Percival, forcing Percival to spend some time determining who to trust, and possibly leading Percival to sabotage the good team by trusting Morgana instead of Merlin.

Good players who are neither Merlin nor Percival (generic good) know nothing about any player other than themselves. Merlin and evil players know the full evil team. Percival knows the two people who are Merlin and Morgana, but not which is which. This selective information is revealed to the appropriate players at the start of the game - termed the "nighttime phase" - by asking players to close their eyes, and then open their eyes to get information or raise their thumb to identify themselves, as appropriate.

In this paper, we seek to evaluate how rules and game size affect the funness of the gaming experience, and how different strategies are more or less successful for good and evil.

2 GAME ANALYSIS

2.1 Evaluating funness

2.1.1 Ideal win ratio. One feature of Avalon is that, not only are the good and evil teams asymmetrical (having different abilities, information, and objectives), but they are also of asymmetric sizes (the evil team in Mafia-style games needing to be a minority to avoid the game being trivially easy).

[1] Avalon may also be played without any special roles (in which case it resembles its predecessor, The Resistance). However, the main feature of Avalon are these special roles.

SIGBOVIK'17, Pittsburgh, PA USA

DOI: 10.1145/nnnnnnn.nnnnnnn

Table 1: Ideal win ratios for each possible size of Avalon game under Supposition B (making evil wins as likely as good wins).

Game Size	# Evil	Ideal Good Win Chance
5	2	.40
6	2	.33
7	3	.43
8	3	.38
9	3	.33
10	4	.40

Supposition A: Funness is maximized by maximizing uncertainty. In other words, we want the probability of winning to be 1/2 regardless which team one is on. This means overall win chance of good and evil must be balanced to be 1/2 for an ideal game-playing experience.

Supposition B: Funness is maximized by making each person's wins equally likely to be earned while on the good team as while on the evil team. Equivalently, across all games, the good and evil teams both produce roughly the same number of winners. Under this supposition:

$$Pr(i \text{ is good} \mid i \text{ wins}) = Pr(i \text{ is evil} \mid i \text{ wins}) = \frac{1}{2} \tag{1}$$

Or...

$$\frac{Pr(i \text{ is good } \cap \ i \text{ wins})}{Pr(i \text{ wins})} = \frac{Pr(i \text{ is evil } \cap \ i \text{ wins})}{Pr(i \text{ wins})} \tag{2}$$

Assuming player i doesn't affect the win probability of their team, this is the same as:

$$Pr(i \text{ is good}) \cdot Pr(\text{good wins}) = Pr(i \text{ is evil}) \cdot Pr(\text{evil wins}) \tag{3}$$

If p_{good} is the win probability for good, and G is the chance of being good (i.e. the number of good roles over the total game size), then this equation is:

$$G \cdot p_{good} = (1 - G) \cdot (1 - p_{good}) \tag{4}$$

$$G \cdot p_{good} = 1 - G - p_{good} + G \cdot p_{good} \tag{5}$$

$$p_{good} = 1 - G \tag{6}$$

In other words, under this supposition of equalizing the portion of good and evil wins, the ideal win ratio for win is $1 - G$, or the chance of being evil.

The number of evil players changes depending on the total number of players. Table 1 summarizes the ideal win chances under this second supposition.

2.1.2 Game duration. Another important component of fun is the length of a board game session. We hypothesize that length of game goes up as the number of players in the game increases due to more discussion. We also are interested in comparing the typical game length to the 30 minutes claimed on the box.

2.2 Strategy

2.2.1 Percival claims. The rules in the instruction manual are not clear on whether players are allowed to claim to be Percival. However, the role of Percival is similar to the role of generic good – and unlike Merlin or the evil roles – in that claiming the role can potentially help the claimant's team (whether good or evil). Therefore, we allow players to publicly claim to be Percival.

A true Percival claim (Percival claiming Percival) can increase trust among good members but can possibly make Merlin assassination easier for the evil team. We are interested in whether claiming to be Percival, and the timing of such claims, tends to help good or evil.

2.2.2 The first mission fail. Evil players have the choice whether to throw in a fail card or a success card for missions that they go on. A sole evil player on the first mission may decide to pass the mission to avoid detection / suspicion for being on a failing team. However, an early mission fail can make the evil team's task of failing three missions total easier.

We are interested in whether first mission fails overall help the good or evil team, and how the size of the first mission factors in to this.

2.2.3 Evil coordination failures. When two or more evil players are on a mission team, they each have to choose whether to throw in a fail or success card, not knowing what their teammates are planning to do. As a result, sometimes evil players may end up passing the mission, or may put in more than one fail card, revealing key information about the make-up of the team. As such, a team with more than one evil person is not ideal for the evil team, and they may be cautious about proposing or approving teams with this make-up.

How often do coordination failures happen, and how do they affect evil's chance of winning? We investigate these questions in this paper.

3 METHOD

A body of 38 graduate students played games of Avalon (20 of which played "semi-regularly" i.e. five or more times during the data collection period). In total, 66 games were recorded although 5 were discarded due to incomplete information, resulting in 61 games overall.

All games were played using the Merlin, Percival, Morgana, and Assassin special roles. In addition, 10 of these games added the special roles Mordred and/or Oberon.

The following data were collected for each game:

- Number of players and role of each
- Approved mission teams and mission outcomes (but not proposed mission teams)
- Outcome of each mission (pass/fail)
- Outcome of game (win for either good or evil), including the win condition (three mission fails (evil win), mission success but Merlin assassination (evil win), or mission success and failed Merlin assassination (only good win condition))
- Which player(s) claimed to be Percival and when (if applicable)

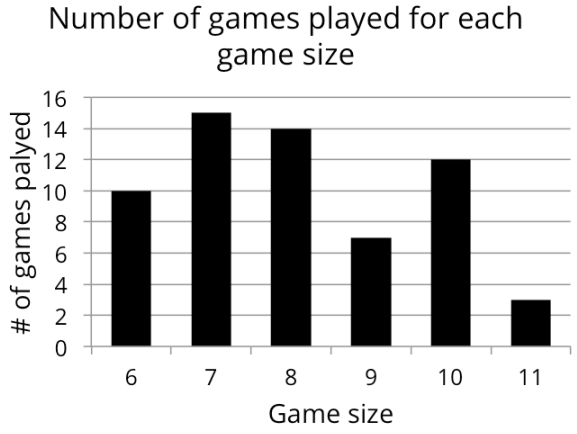

Figure 1: Number of games played for each size of game.

- Which player was assassinated (if applicable)
- Duration of game (measured from the end of night-time phase to either three mission fails or evil's Merlin assassination choice)

Linear regression was used to determine whether game size affected duration.

Chi-squared tests were used to determine whether (1) the presence of Percival claims affected the evil's team Merlin guess rate (Percival claim/no claim vs. Merlin guess/good win condition); (2) evil failing the first mission affected the chance of evil winning (first mission pass/fail vs. winner of game); (3) the presence of missions requiring evil coordination affected the chance of evil winning (zero/non-zero coordination missions in game vs. winner of game).

4 RESULTS

4.1 Number and size of games

Fig. 1 shows how many games of each size were played in the dataset. Although Avalon can in theory be played with game sizes of 5 to 10, players did not enjoy games of size 5 and so only played Avalon if at least 6 players were present. A game of size 11 can be played with the 10-player board and 4 evil characters (as in a 10-player game), and an extra set of vote tokens.

4.2 Win ratio

Overall, the win rate of the good team was .34. Fig. 2 shows how this win ratio changes with the size of the game. The good win ratio for 9-player games stands out as unusually high. This is also the game size with the fewest data points (see Fig. 1), so that may be part of the reason.

Under Supposition A of game funness, 6- and 9-player games are the only ones close to the ideal difficulty for good. 7-, 8-, and 10-player games fall short of both ideal win ratios. As such, it may be worthwhile to use gameplay mechanics that tilt the game in favor of good (Oberon as one of the evil roles, Lady of the Lake, etc.).

Under Supposition B of game funness, the 6- and 9-player games need to be altered to be more difficult. In particular, a 9-player game might benefit from 4 evil roles (instead of 3), one of which is Oberon.

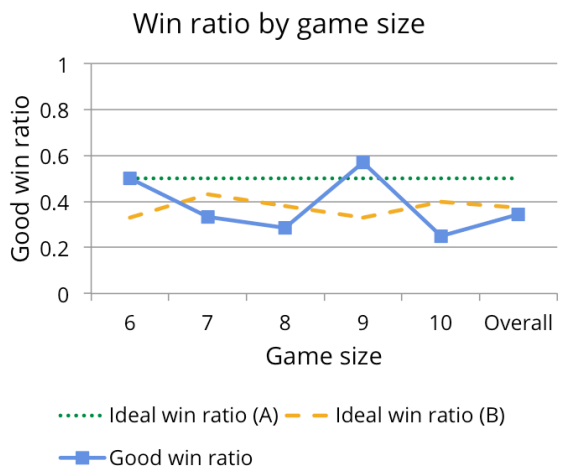

Figure 2: Good's win record at each game size. The dotted line shows the ideal .5 win ratio under the supposition that good and evil should be equally likely to win. The dashed line shows the ideal win ratio under the supposition that people earn wins equally as good people as they do as evil people.

4.3 Game duration

Fig. 3 shows the distribution of game length. The mean game length is 57.3 minutes and the median game length is similar – 57 minutes. Most (80% of) games can be played within 80 minutes. This is markedly longer than the 30 minutes estimated in marketing materials.[2]

We can break down game length by the size of game, resulting in Fig. 4. Games of size 9 are again an outlier, being unusually quick, and being the only game size that approaches the 30-minute estimated play time.

There seems to a slight trend of longer games with more players in line with our hypothesis; however linear regression (removing the 9-person games) does not quite reach significance ($F(1,45)=3.54$, $p=.0663$) and game size has low explanatory power for duration ($R^2=.0524$).

4.4 Percival claims

Fig. 5 compares the outcome of games where a Percival claim is made vs. ones where no Percival claims are made. Games with Percival claims are much more likely to end in mission failure, which makes sense because one reason why Percival might claim is because several failing missions have happened, and Percival (or Merlin) is trying to increase the chance of choosing an all-good team (i.e. scenarios with multiple failing missions are scenarios where Percival is likely to claim).

[2]It is possible that this particular group of graduate students discusses an unusually large amount during Avalon.

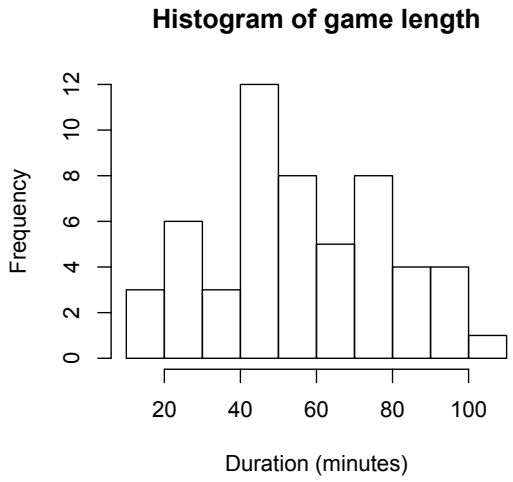

Figure 3: Histogram of game lengths.

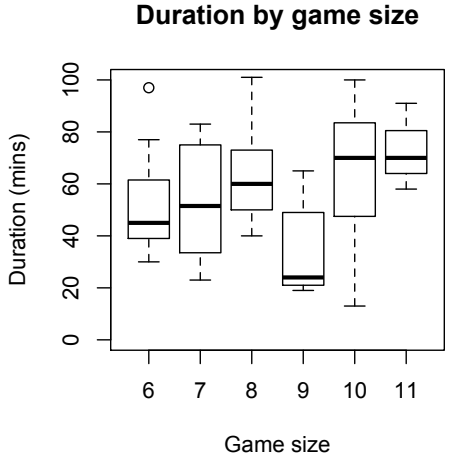

Figure 4: How length of game changes with game size.

Among the remaining cases where three missions succeed, we are interested in comparing how often Merlin is assassinated in each case (Percival reveal vs. no reveal). Table 2 shows the number of games in each condition. The Merlin assassination chance when Percival reveals (.59) is higher than when there is no Percival reveal (.50). However, the chi-squared test shows no significant difference (χ^2=.11, p=.738).

There was not enough data to do an analysis of how the timing of Percival claims affect good's chance of victory. We leave this to future work.

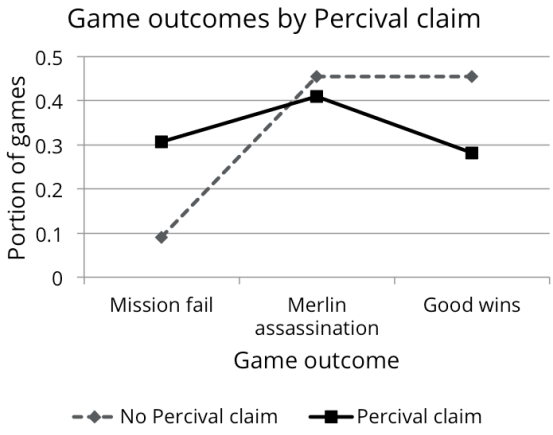

Figure 5: Portion of games that end with mission fails (evil win), Merlin assassinations (evil win), or neither (good win) under the conditions of Percival not revealing and Percival revealing.

Table 2: Merlin assassination successes under the conditions of a Percival claim vs. no such claim.

	Merlin assassinated	Good wins
Percival claim	16	11
No Percival claim	10	10

Table 3: Game victor under the conditions of a first mission fail vs. a first mission pass (games sizes 8+).

	Evil	Good
First mission fail	13	3
First mission pass	5	4

4.5 The first mission fail

Although the intent was to analyze how first mission team size (two-person first missions (in games with 5-7 players) vs. three-person first missions (in games with 8+ players)) affects the outcome of the game, in practice, only one two-person first mission with an evil player (out of 16) was failed by the evil player. This is a difficult strategy to pull off for the evil player because for the rest of the game, at least one good person knows for sure one member of the evil team, and the evil person must consistently behave to give the impression of being someone in that situation.

Therefore, we instead look only at games with three-person first missions. Of the 24 games with evil players present on the first mission, 16 (67%) were failed by those players. Table 3 shows how evil's play during the first mission affected the victor of the game.

There is not enough data to perform a reliable chi-squared analysis, but it is possible that failing the first mission is overall a good strategy for the evil team.

Figure 6: Number of games featuring zero, one, two, or three missions with more evil people than the required number of fails.

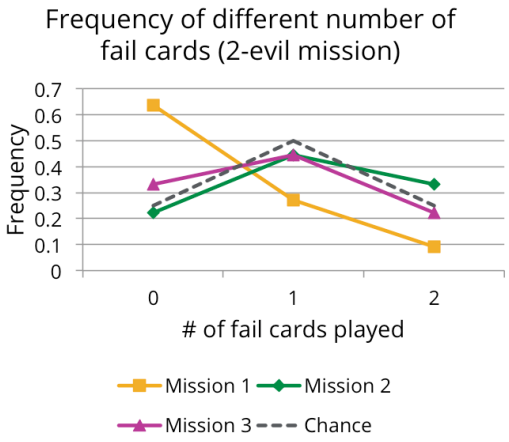

Figure 7: How often zero, one, or two fails come out for one-fail-required missions.

4.6 Evil coordination failures

Of the 61 games, 32 (52%) featured no evil coordination missions, while the rest had at least one evil coordination team. Fig. 6 indicates how often games featured a certain number of evil coordination teams. Overall, this suggests the chance of any mission containing multiple evil people is roughly 15%. Note: It is impossible to have coordination issues on the fifth mission because any number of fails is acceptable for the evil team. Coordination issues on the fourth mission of games of 7+ players, where two fails are required, are rarer (as this only happens when three evil people are placed on the team) but are still important to the game.

38 of the 41 coordination missions (92%) involved two evil people on one-fail-required missions. Fig. 7 summarizes how frequently zero, one, or two fails come out in this situation. This figure shows that the number of fails that come out in Mission 2 and Mission 3 are roughly what you'd expect based on random chance (independent events with .5 probability of occurring). However, Mission 1 is much more skewed toward zero fails, corresponding to roughly a 20-25% chance of each person throwing out a fail. This makes sense as a two-fail result on the first mission can be costly to the evil team.

We also analyze how evil coordination missions affect the chance of good or evil winning the game. Removing from consideration games where evil never gets the chance to go on any mission, Table 4 summarizes the game outcomes when there are no evil coordination missions vs. when there's at least one. Evil's win ratio in the presence of coordination missions (.76) is higher than when there are no coordination missions (.60), although this effect does not reach significance ($\chi^2(1)$=.92, p=.338).

Figs. 8 and 9 break down the effect of evil coordination further. Fig. 8 takes into account evil's coordination performance – success means throwing out exactly the number of fails needed to fail the mission, and failure means throwing out more or fewer fails than needed. Fig. 8 separates the data into three conditions: games where evil mostly failed at coordinating (14 games), games where evil both succeeded and failed at coordinating once (4 games), and games

Table 4: Game victor under the conditions of zero or at least one evil coordination mission.

	Evil	Good
No evil coordination missions	15	10
1+ evil coordination mission	22	7

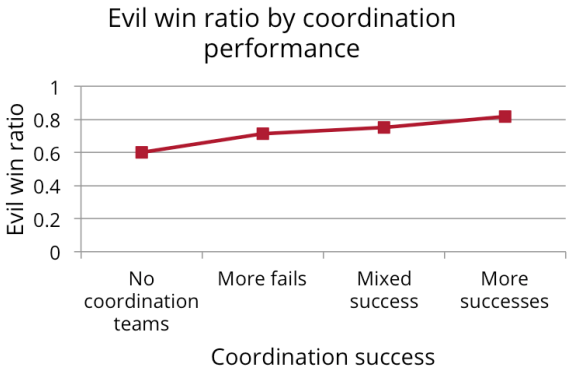

Figure 8: Evil win rate broken down by ability of evil to coordinate.

where evil more often succeeded at coordinating (11 games). Fig. 9 shows how evil's win rate changed as the number of coordination missions in the game increased.

In all cases, there is not enough data to draw any definitive conclusions. However, contrary to expectations, it is possible that evil coordination situations might be slightly beneficial to the evil team.

5 DISCUSSION

We analyzed data from 61 games of Avalon. We found that games of size 9 were unusual in the amount they were played (less popular),

Figure 9: Evil win rate broken down by how many times during the game evil needed to coordinate.

how long they lasted (shorter), and game difficulty (good team more likely to win). Some games might require adjustment to their difficulty. In particular, 9-player games might require more evil characters, and 7-, 8-, and 10-player games might require a slight good handicap. Typical game time is more than one hour.

In this particular dataset, Percival reveals resulted in slightly more Merlin assassinations; evil failing the first mission resulted in more evil wins; and the presence of evil coordination missions resulted in more evil wins. One possible explanation for this last finding is that it is hard for good people to reason about teams where more than one of the members was evil, and so they may be more likely to make decisions that assume only one evil person was on the team.[3] However, more data might reveal these trends to be spurious/random noise.

A notable gap in the dataset is the general absence of two-person first missions failed by an evil player on the team. In the future, it would be interesting for evil to experiment with failing two-person missions to see if this strategy might be beneficial for evil overall. A more detailed analysis of Percival claim timing would be also be good to do with more data. Another promising avenue of future work would be to analyze the effect of rejecting missions on good and evil's chance of winning.

5.1 Conclusion

Hidden role games like Avalon provide a large space for both game design (e.g. number of players, set of specialized roles) and player strategy (e.g. failing the first mission, claiming Percival, team approval strategies, etc.). As such, collecting and analyzing data under different game conditions can be useful in improving the player experience and evaluating the strength of different strategies. Although limited by the amount of data, this work represents a preliminary step in the direction of analyzing the gameplay of Avalon.

ACKNOWLEDGMENTS

The author would like to thank Kristy Gardner for kick-starting this collective Avalon Problem in the department, Ryan Kavanagh for taking better care of the Avalon stats sheet than the author, and Liam K. Bright for listening to the author's shaky probability math.

[3]If evil does indeed benefit from coordination issues, this might interestingly increase the value of Zhu-Brown strategies (multiple evil proposing multiple evil people on their teams).

REFERENCES

[1] Don Eskridge. 2012. *The Resistance: Avalon.* Indie Boards and Cards. Board game.