Counterfeit Coins

Student Researcher: Vasisht Perivilli at Poolesville High School

Mentored by Dr. William Gasarch at The University of Maryland

Overview & Game Mechanics

In this game, a WeighterAgent aims to locate and determine what a set of counterfeit coins is in a set of coins. The counterfeit coins are either heavy or light compared to the normal coins.

Consider a group of N coins. Within this collection, there are H heavy counterfeit coins and L light counterfeit coins, with the remaining coins being genuine. The total number of initial possibilities, representing every valid way to place these counterfeit coins among the set, is determined by calculating the number of possible combinations:

We are given the question: given n coins, and we are told that L are counterfeit and are lighter than real coins,

H are counterfeit and are heavier than real coins, determine in an optimal number of weighings, which coins are counterfeit. We view this as a game that AI will play involving the two AI players, Weighter and Spoiler.

It should be noted that the game of one counterfeit coin has been solved by humans, although the game of one heavy coin and one light coin has not. This is the case of the most interest to us.

Total Possibilities = nCh * (n-h)Cl

To narrow down the suspect coins, we use a two-pan balance scale. On each turn, an equal number of coins is placed on the left pan and the right pan. The scale provides one of three clear physical outcomes:

Left Side Heavy: The coins on the left pan outweigh the coins on the right pan.

Balanced: Both pans weigh the exact same amount. This occurs when both pans contain no counterfeit coins, equal numbers of identical fake coins, or when heavy and light coins on the same pan cancel out each other's weight differences.

Right Side Heavy: The coins on the right pan outweigh the coins on the left pan.

Since there are three different possible weighings, the theoretical lower bound which was tested to be $\log_3$(nCh *

(n-h)Cl).

Agent Protocols:

This experiment was completed with two agents: The Spoiler and the Weighter. Both of them are described in this section.

The Spoiler Protocol

Search performance was tested against a SpoilerAgent.

Under this framework, no specific coins are designated as fake at the start of the game. This is vital to our understanding: Neither the weighter nor the spoiler actually know which coins are counterfeit, just how many are light and how many are heavy. Instead, all possible coin arrangements remain active in a pool of suspect configurations. Whenever the weighter places coins on the scale, the prospective move splits the suspect pool into three groups based on what the scale would read for each possible arrangement.

The Spoiler evaluates the size of these three resulting groups and actively chooses the outcome that leaves behind the largest remaining group of suspect arrangements. The Spoiler then declares that outcome as the scale result. Crucially, the Spoiler must remain consistent in that every result it gives must agree with all prior scale readings. The game continues until the suspect pool is reduced to exactly one remaining possibility, forcing a single final arrangement that satisfies every weighing performed during the game.

This is regarding the first agent, the SpoilerAgent. The WeighterAgent, which performs the weighings themselves, is described in the next section.

The Weighter Protocol

Because the Spoiler always selects the largest outcome group to prolong the game, the weighter cannot rely on lucky guesses. To minimize the total number of weighings under worst-case conditions, the weighter must rely on the following procedure.

Since the balance scale produces three possible outcomes, the most efficient move for the WeighterAgent is one that divides the current pool of suspect arrangements into three equal parts (approximately 33.3% of the suspect

pool per group).

If a weighing produces unbalanced groups, for example, leaving 80% of possibilities in the "Left Heavy" group, 10% in "Balanced", and 10% in "Right Heavy", the Spoiler will intentionally select the "Left Heavy" reading. This leaves 80% of the suspect pool alive, making minimal progress. By contrast, if a move divides the suspect pool evenly into three 33.3% groups, the Spoiler is forced to choose a group containing only 33.3% of the possibilities, eliminating two-thirds of the remaining suspect arrangements in a single turn.

This is the objective of the agent. However, this almost never happens, and there is no way to predict when we can get a perfectly even split. This will lead to some data being over the theoretical lower bound at some points. However, these are not very common, and the weighter agent does perform well in general.

Findings & Detailed Analysis

The Weighter was tested against the Spoiler across several coin counts and counterfeit structures. Since the spoiler consistently chooses the largest remaining suspect group and the Weighter operates under a specific set of rules, every trial for a given setup requires the exact same number of weighings.

Amount of Coins	**Distribution of Counterfeit Coins**	**Total Possibilities: nCh*(n-h)Cl**	**Estimated Amount of Weighings ($\log_3$[Total Possibilities] rounded up)**	**Actual Amount of Weighings**
8 Coins	1 Heavy, 1 Light	56	4 weighings	**4**
10 Coins	1 Heavy, 1 Light	90	5 weighings	**5**
12 Coins	1 Heavy, 1 Light	132	5 weighings	**6**
14 coins	1 Heavy, 1 light	182	5 weighings	**6**
16 coins	1 Heavy, 1 light	240	5 weighings	**6**
18 coins	1 Heavy, 1 light	306	6 weighings	**6**
20 coins	1 Heavy, 1 light	380	6 weighings	**6**
22 coins	1 Heavy, 1 light	462	6 weighings	**6**
12 Coins	2 Counterfeits (All heavy)	66	4 weighings	**4**
8 Coins	3 Counterfeits (All Heavy)	56	4 weighings	**4**

Unfortunately, more data could not be collected due to computational limits.

Situation 1: 12 Coins with 2 Counterfeits (Known to be All Heavy)

In this scenario, there are 12 total coins and exactly 2 counterfeit coins, both known in advance to be heavy. The initial suspect pool consists of 66 possible pairs.

Because both fake coins are heavy, placing them on opposite pans creates a clear tilt toward whichever pan holds a fake coin. Placing both on the same pan simply doubles the downward pull. Crucially, two fake coins can never offset each other's weight on the scale. This removes ambiguous readings, allowing the weighter to split the 66 suspect arrangements cleanly across consecutive weighings. The Spoiler cannot maintain a large candidate group, forcing the game to finish in 4 weighings.

Situation 3: 8 Coins with 3 Counterfeits (Known to be All Heavy)

In this scenario, there are 8 coins containing 3 counterfeit pieces, all known to be heavy. The initial suspect pool contains 56 possible combinations.

With 3 heavy coins in circulation, testing equal groups of coins on the scale produces distinct weight differences (such as one side having 1 extra heavy coin, 2 extra heavy coins, or all 3). These strong, unmistakable scale deflections allow the weighter to partition the suspect pool effectively. The Spoiler is unable to preserve large suspect groups across turns, resolving the search in 4 weighings.

Why the 12-Coin, 14-Coin, and 16-Coin Mixed Setups Take 6 Weighings

For 12 coins with 1 heavy and 1 light coin (132 initial possibilities), the theoretical minimum based on tripling information on each move is 5 weighings. However, against the Spoiler, the game always takes 6 weighings.

The weighings are as follows:

- [0, 1, 2, 3], [4, 5, 6, 7]: 48, 42, 48
- [0, 1, 8, 9], [2, 4, 5, 10]: 18, 15, 15
- [0, 2], [1, 8]: 7, 6, 5
- [0,3], [4, 9]: 3, 2, 2: Candidates: 0, 9; 0, 4; 3, 9
- [4], [10]: 0, 2, : Candidates: 0, 9; 3, 9

- [0], [3]: 1, 1: Candidates: 0, 9

This list is in the format: [Left pan], [Right pan]: Amount of left-heavy configurations, balanced configurations, right heavy configurations, candidates for the heavy and light coins.

We see here that from 7, which is less than 3^2, we go to 3 in the 5th weighing. This is still only less than 3^2, meaning that an extra weighing must be added to the theoretical lower bound, resulting in 6 weighings.

All other situations end up in the theoretical lower bound, because we do not face this issue.

Although the lower bound is not a guaranteed amount of weighings, we can absolutely treat it as a minimum amount of weighings. Despite this, there is no specific proof for any value of n coins that a certain amount of coins will take more than the lower bound, and we almost have to leave it to chance.

Improvements and Conjectures:

The biggest limitation in our data is that we cannot get data for any amount of coins greater than 22. This makes our data limited, but based on the data that we have, we can make some inferences and conjectures. We can assume that the upper bound for the amount of weighings would be the lower bound ($\log_3$(nCh * (n-h)Cl)) + 1. This simplifies to $2\log_3 n$ for large n, meaning we can assume that for large values of n, the upper bound would be $2\log_3 n$.

Acknowledgements

Thank you to Dr. William Gasarch at the University of Maryland for mentoring this work during the summer of 2026, and to Isaac Yoo, Himanshu Deshpande, Hargun Malhotra, and Sidanth Poduri for working together with me through this research and giving me advice throughout. Some code was debugged using generative AI.