Solving Wordle Heuristically *(DRAFT — February 6, 2022)*

Daniel W. Dichter
Independent Researcher
Cambridge, Massachusetts, U.S.A.
daniel.w.dichter@gmail.com

Wordle is a daily word game in which a player must guess a hidden five-letter word in six or fewer tries. After each guess, they get a color-coded response that indicates how close their guess was. This paper describes a solver that wins all 2,315 possible games, averaging a score of 3.692 with a standard deviation of 0.734. This is comparable to the current best solvers known, yet is achieved heuristically with many fewer calculations.

I. Game Parameters

Wordle has two dictionaries: 2,315 possible solutions, and 12,972 acceptable guesses.[1][2] Solutions are limited to commonly-known, non-S-plural words. The default version of the game is assumed, wherein guesses need not be consistent with known information. Games are assumed to be mutually independent, and thereby the solver cannot recall previous solutions to reduce the set of initial words.

II. Eliminating Solutions

After each guess, some solutions can be eliminated. This is executed sequentially, tile-by-tile, as follows.

Gray tile. If any and all additional letter-matching tiles are also gray, eliminate words that contain this letter. Otherwise, eliminate words that contain this letter in this position, and eliminate words that do not contain the overall quantity of this letter indicated.

Yellow tile. Eliminate words that contain this letter in this position, and eliminate words that do not contain this letter in non-green non-letter-matching tiles.

Green tile. Eliminate words that do not contain this letter in this position.

Game State	Words Left
S O A R E	2,315
F L I N T	63
A P E A K	3
S K I L L	1
S O A R E	2,315
A L O F T	44
S O A R E	2,315
P E T A L	68
P L E A T	1
S O A R E	2,315
T H A C K	11
S H A R D	2

Table 1. Solutions for recent games.

III. Probability and Value Matrices

At any game state, a matrix measuring 26 x 5 may be defined that reports the probability of each letter in each position of the solution.

Guessing a tile with zero or unity probability[1] provides no information. This observation is generalized by stating that "bad" guesses have high certainty. It follows that "good" guesses have low certainty, which is equivalent to intermediate probability. This is formalized as transfer function relating probability P to value V, modeled as one period of a sine function with $V(P = 0, 1) = 0$ and $V(P = {}^1\!/_2) = 1$. Gaussian and triangular functions intersecting these same (P, V) points with equal $\text{sign}(dP/dV)$ were found to perform marginally worse.

Another matrix, also measuring 26 x 5, may then be defined that reports the value of each letter in each position of the solution.

IV. Scoring and Guessing

After every guess, the value matrix is updated per the remaining solutions and used to score all guesses. Scoring is also executed sequentially, tile-by-tile. Interactions between tiles are modeled with a set of heuristics, as follows.

A tile with non-unity probability captures the value of that letter in that position, and one-half the value of that letter in all other positions; with unity probability it captures no value in any position. Guesses with multiple tiles of the same letter interfere with each other; their value capture is equal to their nominal sum divided by their quantity, excepting letter-matching unity-probability tiles. Guesses where the quantity of letter-matching tiles is less than or equal to the quantity of unity-probability letter-matching tiles may refer back to the unity-probability tile(s) as yellow; tiles matching this condition are diminished in value by a factor of one-half.

The top-scoring word is picked until there are two or fewer solutions remaining. Ties are broken first by solution probability, then alphabetically.

---

[1] Obtained directly by a gray or green tile, or indirectly by deduction.

V. Results

Top-scoring first words are shown in Table 2. Solutions for recent games are shown in Table 1. Results for all games are shown in Table 3. A comparison with other solvers including unassisted play by proxy of Twitter[2] is shown in Table 4 and Figure 1.[3] Source code and complete game logs are also provided.[4]

VI. Discussion

All games are opened with the top-scoring word SOARE. In the middlegame, emergent strategies can be seen, such as abandoning green tiles until the endgame, shifting yellow tiles to alternate positions, and reusing gray tiles for high-value words. The cumulative density function (CDF) shows that 88.5% of games are won within four moves, 99.0% within five, and 100% within six. The probability density function (PDF) shows a mode of four moves, with an associated probability of 47.9%. The average score is within 7% of the current best solvers known, which are substantially more complex and/or computationally expensive. An average score lower limit of about 3.5 is theorized based on the maximum amount of information that may be obtained in the first two moves.[5]

VII. Acknowledgments

The author gratefully thanks Jess Hall, Julius Dichter, Adam Filion, and Tyler Smith for their support.

Rank	Word	Score
1	SOARE	1.00000
2	STARE	0.93367
3	SAINE	0.92580
4	SAICE	0.91265
5	SLATE	0.91234
6	RAISE	0.91178
7	AROSE	0.91162
8	SNARE	0.90631
9	ARISE	0.90500
10	STOAE	0.90352

Table 2. Top-scoring first words.

Score	Games	PDF	CDF
1	0	0.000	0.000
2	63	0.027	0.027
3	876	0.378	0.406
4	1,110	0.479	0.885
5	242	0.105	0.990
6	24	0.010	1.000
x	0	0.000	1.000

Table 3. Results for all games.
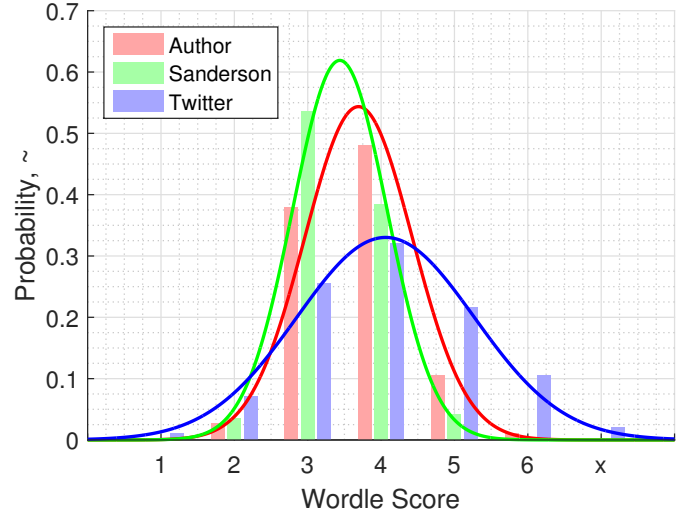
Figure 1. Comparison of solver distributions.

Solver	Win %	Mean	Stdev.	Worst	Opener
Author[4]	100.0	3.692	0.734	6	SOARE
Sanderson[5]	100.0	3.438	0.645	6	CRANE
Glaiel[2]	100.0	3.494	-	5	ROATE
Chao[6]	99.7	3.420	-	x	OPERA
Filion[7]	94.2	-	-	x	AROSE
Twitter[3]	97.9	4.061	1.208	x	–

Table 4. Comparison of solver benchmarks.

References

[1] Wardle, Josh. *Wordle - a daily word game.* Accessed February 6, 2022. https://www.powerlanguage.co.uk/wordle/.

[2] Glaiel, Tyler. *The mathematically optimal first guess in Wordle.* Accessed February 6, 2022. https://medium.com/@tglaiel/the-mathematically-optimal-first-guess-in-wordle-cbcb03c19b0a.

[3] O'Connor, Kevin. *Wordle Stats (@WordleStats) / Twitter.* Accessed February 6, 2022. https://twitter.com/wordlestats.

[4] Dichter, Daniel. *wordle_solver.* Accessed February 6, 2022. https://github.com/kindofdoon/wordle_solver.

[5] Sanderson, Grant. *The mathematically optimal Wordle strategy.* 3Blue1Brown. Accessed February 6, 2022. https://www.youtube.com/watch?v=v68zYyaEmEA.

[6] Chao, Jason. *Wordle-Solver.* Accessed February 6, 2022. https://github.com/jason-chao/wordle-solver.

[7] Filion, Adam. *Building a Wordle Solver.* Accessed February 6, 2022. https://blogs.mathworks.com/loren/2022/01/18/building-a-wordle-solver/.