Entropy-Driven Strategies in Guessing Games

Group A

June 4, 2025

Introduction

In information theory, entropy H quantifies the average uncertainty of a random variable. Word guessing puzzles such as Wordle or Mastermind can be modelled in exactly this way before any clues are revealed, the hidden word W is assumed to be uniformly distributed over a dictionary aftereach guess, the feedback partitions the dictionary and thereby reduces H(W) . The expected drop in entropy, given the information gain, offers a principled metric for choosing the next guess.

Matteo Da Ros's thesis demonstrates the idea in the full 12972 word Wordle dictionary. For every admissible guess w, he enumerates the 243 possible feedback patterns, computes the entropy X_w of that feedback random variable, and plays the guess that X_w maximises X_w. The resulting solver finishes the game in an average of 4.1 guesses, matching strong human play and outperforming naive frequency strategies.

The concepts of entropy and information

In probability theory, the concept of information is closely related to the amount of surprise that an event carries with itself. This means that an event with high probability of happening carries less information than an event with a lower probability. A good candidate to compute and quantify this concept is the following:

$$\mathrm{I(E)=\text{-}K}\ \log_a(P(E))$$

In our computations we will consider K=1 and a=2. The unit of measure of information is the bit. Consider, for example, a Bernoulli random variable with p=0.5:$X \sim Ber(0.5)$ and the events X=1, X=0. The information of these events is:

$$\text{I(X=1)=I(X=0)=-} \log_2(0.5) = 1 \text{ bit}$$

So both these events carry exactly 1 bit of information, and this explains why the choice of K=1 and a=2 is convenient.

Let's now consider the discrete random variable X defined above. We wish to compute the amount of information expected from the execution of the experiment. We can compute this by calculating the expected value of the information of all possible outcomes.
This quantity is called entropy and denoted by H(X):

$$\text{H(X)=E[I(X)]=P(X=1)*I(X=1)+P(X=0)*I(X=0)=0.5+0.5=1}$$

which is not so exciting, since the two events carry the same amount of information. Let's now consider a Bernulli random variable with generic probability p. Then its entropy is:

$$\text{H(X)=-p}\log_2(p) - (1-p)\log_2(1-p)$$

Let's compute this quantity for different values of p:

$$\text{H(Ber(0.75))=0.811}$$

$$\text{H(Ber(0.9))=0.469}$$

It can be proven that the entropy is maximum when there is a uniform distribution. The general formula of entropy, considering a random variable X with probability law p0,...,pnis: The entropy $H(X)$ is given by:

$$H(X) = -\sum_{t=0}^{n} p_t \log_2 p_t$$

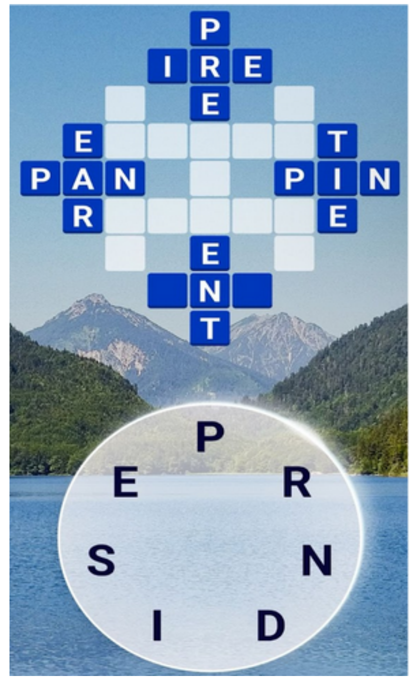
Entropy is a measure of uncertainty: the higher it is, the less information we have about the random variable, since entropy represents the expected information that we will get from the outcome of the experiment. On the other hand, an entropy equal to zero implies that the random variable takes one value with probability 1 and all the other outcomes have probability zero, which means that there is no uncertainty.

By applying this concept to a word guessing game, we can develop a strategy to minimize the entropy after each guess, in order to get to the solution in the most efficient way.

Game: Word Connect

Conceptual Blend:

- Anagram Puzzle: Form words from scrambled letters
- Combinatorial Optimization: Maximize word yield from limited letters
- Vocabulary Challenge: No feedback mechanism - players self-validate



Rules:

– Fixed multiset of letters (e.g., P, I, R, E, E, T, P, A, N)

– Target word length specified per level (e.g., 3+ letters)

Objective: Form all possible valid English words of the target length using:

– Subsets of the available letters

– Letter frequencies constrained by the multiset

– No positional constraints (unlike Wordle)

Constraints: For candidate word w: count(letter in w) $\leq$ frequency(letter in multiset) for all letters $\ell \in A$
Example: from R, T, A, C, B, valid 3-letter words include:

– CAT, RAT, BAT, TAB

– But not EAT (E unavailable)

Gameplay: No feedback mechanism Players submit solutions when they believe all words are found Victory requires complete enumeration of valid words

Entropy Computation: Initial Entropy: Set N as the number of total valid words

$$H_0 = N \cdot \log_2(N)$$

Solution Space Represents player's uncertainty about the complete solution set

Example: For R, T, A, C, B at L = 3: Valid words = CAT, RAT, BAT, TAB, ACT, ART, BAR, CAR, RAB

$$H_0 = 9 \cdot \log_2(9) \approx 28.53 \text{ bits}$$

Progress Metric - Entropy reduction occurs only upon complete solution - Final entropy: H$x_i$final = 0 (complete certainty)

Graphical Representation: Valid solutions: 4-letter: PINE, RENT, PANE, TIRE, PEAR 3-letter: PIN, TAP, NAP, RAT, PAT, NET) (Note: This is an illustrative example)

– High initial H = Many possible valid words

– $H \propto$ vocabulary size and multiset flexibility

– Zero feedback maintains constant entropy during gameplay

Uncertainty dynamics: This model represents a combinatorial entropy minimization problem where:

– Entropy quantifies the player's uncertainty about the full solution set

– Complete knowledge = Exhaustive enumeration of valid words

– Game difficulty $\propto H_0$ (higher entropy = harder levels)

Entropy-Based Guessing Strategy

Since each word in (V) has a uniform probability of being in (S) (p = 5/11), and within (S), each word has an equal probability(1/5)each guess provides the same amount of information. The strategy is to pick a random word from (V), check if it's in (S), and track the entropy after each guess to measure the reduction in uncertainty.

Setup: Multiset: L = Y,R,S,N,T,A
Valid words ((V)): ANY, ART, ANT, RAT, RAN, RAY, SAT, SAY, TAN, TAR (11 words).
Solution set ((S)): ANY, ANT, RAT, RAN, ART, TRY (5 words, unknown to the player).
Initial entropy over (S) : H0 = log(5)= 2.333 bits.

Random Guessing with Entropy Tracking:

Step 1: Randomly pick a word from (V). Choose ANY.
ANY is in (S). Found = ANY.
Remaining words in (S): ANT, RAT, RAN, ART, TRY (4 words).
Entropy: H1 = $\log_2(4) = 2$ bits.

Step 2: Pick RAY.
RAY is not in (S). No update.
Entropy: H2 = $\log_2(4) = 2$ bits.

Step 3: Pick RAT
RAT is in (S). Found = ANY, RAT.
Remaining: ANT, RAN, ART, TRY (3 words).
Entropy: H3 = $\log_2(3) = 1.585$ bits.

Step 4: Pick SAT.
SAT is not in (S). No update.
Entropy: H4 = $\log_2(3) = 1.585$ bits.

Step 5: Pick ANT.
ANT is in (S). Found = ANY, RAT, ANT.
Remaining: RAN, ART, TRY (2 words).
Entropy: H5 = $\log_2(2) = 1$ bits.

Step 6: Pick RAN.
RAN is in (S). Found = ANY, RAT, ANT, RAN.
Remaining: ART, TRY (1 word).
Entropy: H6 = $\log_2(1) = 0$ bits(assuming the player knows ISI = (5).

Step 7: Pick ART.
ART is in (S). Found = ANY, RAT, ANT, RAN, ART.
Remaining: TRY.

Step 8: Pick TRY.
TRY is in (S). Found = ANY, RAT, ANT, RAN, ART, TRY.
All words in (S) found. Final entropy: Hfinal = 0.

Submit the Solution: Final list: ANY, RAT, ANT, RAN, ART, TRY (matches (S)).

Limitations: Without knowing (—S—), the player may not know when to stop. Random guessing can be inefficient in practice; a heuristic like prioritizing common letter patterns might speed up the process, but wouldn't change the entropy due to uniformity.

Conclusion

Treating a guessing game as an entropy-minimisation process yields two complementary insights:

1. Analysis – entropy tracks how much uncertainty remains after each turn
2. Strategy – maximising expected information gain provides a simple yet effective rule for selecting guesses.

Empirically, the entropy maximising policy on Wordle reduces the initial 12972 candidates to roughly six after just two moves, underscoring the power of information theoretic planning.

Bibliography

M. Da Ros – Un approccio risolutivo per Wordle utilizzando la teoria dell'informazione. Bachelor Thesis, Università di Trento (2021).

C. E. Shannon – "A Mathematical Theory of Communication" Bell System Technical Journal, 1948

Game rules and tips: **https://wordfinder.yourdictionary.com/word-connect-answers/**

Cover, T. M., & Thomas, J. A. (2006). Elements of Information Theory. Wiley.

Carter, T. "An Introduction to Information Theory and Entropy."

Merriam-Webster Dictionary. Accessed for validating English words. Available at:**https://www.merriam-webster.com/**