Fighting the Hydra: A Finite Game that Escapes Arithmetic

Oliver Au

The Kirby–Paris hydra game is inspired by a famous myth from ancient Greece. In the story, the hero Hercules battles a creature called the Lernaean Hydra, which can regrow its heads when they are cut off. At first, this might sound like a simple tale of persistence—but in terms of the mathematical game, it leads to something much deeper.

In the hydra game, we imagine a similar creature made of branching "heads." Each time one is cut off, the hydra grows back in a specific way, sometimes becoming even larger than before. Surprisingly, no matter how the game unfolds, Hercules always eventually wins, defeating the hydra after finitely many steps. However, this result hides an unexpected twist. Although "Hercules always wins" is true, this fact cannot be proved using ordinary first-order arithmetic alone.

In this essay, I will explore why intuitive arguments fail to show that the game must end and introduce a powerful idea—an ordinal-valued measure—that allows us to prove termination as well as how the hydra game links to the Goodstein sequences, another surprising mathematical phenomenon.

A hydra is a finite rooted tree. The root is the body and each leaf is a head. A battle proceeds in stages numbered $n = 1, 2, 3, \ldots$. At stage n Hercules chooses one head and chops it off. If the head was attached directly to the root, then it simply disappears. If it was higher up, the hydra regenerates by duplicating a nearby subtree in a way that depends on n. If Hercules chops a head h, if the predecessor of h is the root, nothing happens; otherwise let h_1 be the father (parent) of h and h_2 the grandfather. Remove the head h, and then sprout n copies of the principal subtree determined by h_1 (with h removed) from h_2.

If we consider a 'strategy' to be some rule Hercules uses to decide which head to cut next, we can find that extraordinarily, every possible sequence of legal chops eventually kills the hydra. Kirby and Paris prove that every strategy is winning, and later expositions often state this as "Every strategy against the Hydra is a winning strategy."

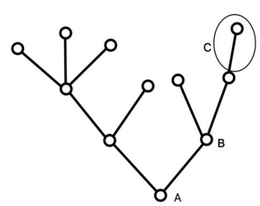


Figure 1: Example of a Hydra.

Now that we have established how to draw the Hydra, it will be easier to describe its regenerating powers:

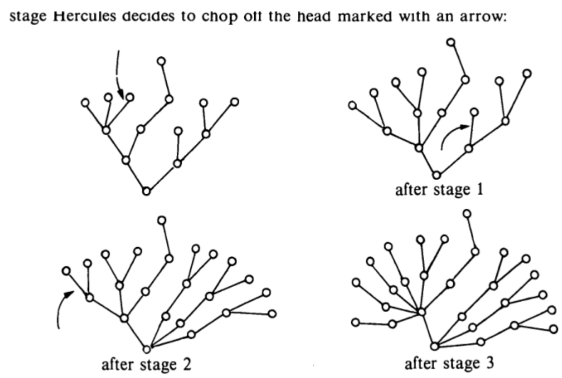
- Hercules for the n^{th} time severs a head.
- Traverse one node down from where the head was severed.
- The entire part of the Hydra from this node up is multiplied n times, and grows from the relevant node.

2 First attempts: why the hydra looks unbeatable

To try and prove that a process terminates, A good first instinct is to find a quantity that decreases each step. Here the most obvious quantity is the number of heads. Let H_n be the number of heads just before the nth chop. In most games one hopes for $H_{n+1} < H_n$. In the hydra game this is false. A single chop can create n copies of a subtree, and each copy may contain several heads. It is therefore typical that

$$H_{n+1} \gg H_n.$$

What about counting edges, total nodes, or height? None of these ieither. This is not just an annoyance; it is the point. The hydra is designed to force you to abandon integer-valued size measures. In real mathematics and computer science this happens often. For example, in some algorithms and rewrite systems a step can increase the size of the data structure even though it moves the computation closer to finishing. The correct proof method is then to assign a more sophisticated ranking into a well-founded order. The hydra is an extreme model of this situation.



3 Ordinals: the right way to measure progress

The standard proof in order to show that all hydra battles terminate assigns to each hydra T somethig called an ordinal $o(T)$ below a particular ordinal threshold ε_0, and then prove that each move strictly decreases this ordinal below this limit. Kirby and Paris who popularised the problem developed exactly this kind of ordinal assignment, working with ordinals $< \varepsilon_0$ and showing that for any strategy a, the ordinal at stage n is obtained from the ordinal at stage $n-1$ by an operation $[a](n)$ that always produces a smaller ordinal.

3.1 What are ordinals?

So first of all to understand this, What is an ordinal? Ordinals generalise counting. The finite ordinals are $0, 1, 2, \ldots$. The first infinite ordinal is ω, representing the order type of the natural numbers. Then $\omega + 1$ means "all naturals, then one more." One can form $2\omega = \omega + \omega$, then ω^2, then ω^ω, and so on. The key property is that there is no infinite strictly decreasing sequence

$$\alpha_0 > \alpha_1 > \alpha_2 > \cdots.$$

This can be seen to be taken from the fact that you cannot count down forever from a natural number, but it remains true for all ordinals.

3.2 Cantor normal form and ε_0

Any ordinal number $\alpha < \varepsilon_0$ can be written in a standard form called *Cantor normal form.* This means we can express α as

$$\alpha = \omega^{\beta_1} c_1 + \omega^{\beta_2} c_2 + \cdots + \omega^{\beta_k} c_k,$$

where the exponents decrease from left to right ($\beta_1 > \beta_2 > \cdots > \beta_k \geq 0$), and each coefficient c_i is a positive whole number.

You can think of this as similar to writing a number in base 10. For example,

$$345 = 3 \cdot 10^2 + 4 \cdot 10^1 + 5 \cdot 10^0.$$

Cantor normal form works in a similar way, but instead of powers of 10, we use powers of ω, which represents an infinite quantity. Just like in base 10, the larger powers dominate the smaller ones, so the term with ω^{β_1} has the greatest influence on the size of the ordinal.

The ordinal ε_0 is a particularly important number. It is defined as the smallest ordinal that satisfies the equation

$$\varepsilon_0 = \omega^{\varepsilon_0}.$$

In other words, it is the first ordinal that remains unchanged when used as an exponent of ω.

In proof theory, ε_0 measures the strength of Peano Arithmetic in a precise sense: it marks the limit of the kinds of arguments that can be carried out within that system. The Kirby–Paris hydra game is carefully constructed so that proving "Hercules always wins" requires reasoning up to this level. This is why the result sits right at the boundary of what Peano Arithmetic can prove, making the hydra game an important example in the study of mathematical independence.

The ordinal ε_0 is the smallest fixed point of exponentiation by ω:

$$\varepsilon_0 = \omega^{\varepsilon_0}.$$

The reason ordinals (up to ε_0) help solve the hydra game is that they provide a way to measure progress, even when the hydra appears to grow larger.

At first, the game seems difficult to control: when Hercules cuts off a head, the hydra may regrow several new heads, so simple measures such as the total number of heads do not decrease. In fact, they can increase.

The key idea is to assign to each hydra an ordinal number $\alpha < \varepsilon_0$ which encodes its structure. This assignment is based on the shape of the hydra as a rooted tree. Branches that are deeper in the tree correspond to higher powers of ω, so the ordinal reflects not just the size of the hydra, but how it is built.

For example, a simple branch might correspond to a term like $\omega^0 = 1$, while more complex branching structures contribute terms like ω^β for larger β. These are combined using Cantor normal form.

When Hercules makes a move, two things happen:

- A head is removed, which decreases the contribution from a higher-level term.
- The hydra regrows new branches, but these only contribute lower-level terms.

Because ordinal arithmetic is dominated by the largest exponent, removing a term like ω^β and replacing it with a finite sum of smaller terms results in an overall decrease. Thus, each move strictly decreases the assigned ordinal.

Since there are no infinite descending sequences of ordinals, this process must eventually terminate. Therefore, after finitely many steps, the hydra is completely destroyed and Hercules wins.

Although the ordinal assignment shows that the hydra game always terminates, there is an important limitation to this argument. The proof relies on the fact that there are no infinite descending sequences of ordinals below ε_0. To justify this rigorously, we must use a strong form of induction called *transfinite induction* up to ε_0.

Peano Arithmetic (PA) is a formal system that captures the usual properties of the natural numbers, including addition, multiplication, and induction. However, the induction proof available in PA applies only to natural numbers, not to the much larger ordinal structures used in this proof.

It turns out that PA by itself cannot prove that all descending sequences of ordinals below ε_0 terminate. Since the hydra argument depends on exactly this fact, the statement "Hercules always wins" cannot be proved within PA, even though it is true.

For this reason, the hydra game is not just an interesting puzzle, but an example of an *independence phenomenon*: a statement that is true, but whose proof requires methods that go beyond the strength of the arithmetic we know.

3.3 Why duplication can still mean descent

Here is an example calculation that illustrates the key mechanism, without requiring the full formal definition of $o(T)$. Suppose a hydra has a "large" component measured by ω^2 and a "medium" component measured by a multiple of ω. Consider the ordinal

$$\alpha = \omega^2 + 3\omega + 5.$$

This expression reflects different levels of structure in the hydra: the ω^2 term represents the most significant part, the ω term represents a secondary level, and the constant term represents smaller details.

Now imagine a move that causes the hydra to grow dramatically, so that the visible number of heads increases a great deal. The ordinal might change to

$$\alpha' = \omega^2 + 2\omega + 10^{10}.$$

Even though 10^{10} is much larger than 5, we still have $\alpha' < \alpha$. This is because the first place the Cantor normal forms differ is in the coefficient of ω, and $2 < 3$. Formally,

$$\omega^2 + 2\omega + 10^{10} < \omega^2 + 3\omega + 5.$$

This illustrates a key principle of ordinal arithmetic: a decrease at a higher level outweighs any increase at lower levels. In other words, "many copies of a lower-level term" are still dominated by "one fewer copy of a higher-level term." This ikind of idea is the same as base-10 arithmetic, where decreasing the thousands digit by 1 outweighs any increase in the units digit.

Kirby and Paris prove that a suitably defined ordinal assignment $o(T)$ satisfies

$$o(T_{n+1}) < o(T_n)$$

for every legal move from T_n to T_{n+1}, regardless of the strategy used. Since there is no infinite descending sequence of ordinals below ε_0, it follows that the process must eventually terminate, and hence Hercules always wins.

A helpful diagram showing several steps of a hydra battle, represented as trees, can be found in Tigerschiöld's exposition, which provides additional intuition for how the process evolves.

4 A parallelism: Goodstein sequences

Hydras are closely related to another process in maths that at first seems as though it should never end, known as a Goodstein sequence. Like the hydra game, it exhibits rapid growth, yet eventually terminates.

A Goodstein sequence starts from an integer $m_0 = m$ and an initial base $b_0 = 2$. The number m_i is written in *hereditary base* b_i, meaning that not only the number itself, but also all exponents in its representation, are expressed in base b_i, recursively. To form the next term m_{i+1}, we replace every occurrence of b_i with $b_i + 1$, and then subtract 1.

A small example illustrates the surprising behaviour. Start with $m_0 = 4$ in base 2. We can write

$$4 = 2^2.$$

Replacing 2 by 3 and subtracting 1 gives

$$m_1 = 3^2 - 1 = 8.$$

Now write 8 in base 3 as

$$8 = 2 \cdot 3 + 2.$$

Replacing 3 by 4 and subtracting 1 gives

$$m_2 = 2 \cdot 4 + 1 = 9.$$

Thus the sequence begins

$$4, 8, 9, \ldots$$

and can grow very quickly, often reaching extremely large values before eventually decreasing to 0. From this the parallel with the hydra game becomes clear

The proof of termination again uses ordinals below ε_0. Although the numerical values increase, one can associate to each term an ordinal that strictly decreases at every step. If the sequence did not terminate, this would produce an infinite strictly decreasing sequence of ordinals below ε_0, which is impossible.

This provides a useful conceptual bridge to the hydra game. In Goodstein's theorem, the numbers increase because the base increases. In the hydra game, the tree can grow because later moves replicate more branches. In both cases, however, the correct measure of progress lies in the transfinite: when translated into ordinals, the process is strictly decreasing, ensuring that it must eventually terminate. Goodstein sequences are important not only because of their surprising behaviour, but also because of what they reveal about the limits of mathematical reasoning. Although every Goodstein sequence eventually terminates, this fact cannot be proved within Peano Arithmetic. As a result, they provide a clear and concrete example of the limitations of formal systems, and play a central role in proof theory.

5 Conclusion

The Kirby–Paris hydra game looks simple at first—almost like a game you could explain to a child. You cut off a head, the hydra grows back more, and you repeat. But very quickly, it challenges your intuition. Counting heads is not enough to understand what is happening, and familiar methods using ordinary numbers begin to fail.

To make sense of the game, we need a new idea: instead of measuring size with integers, we measure progress using ordinals, which extend beyond finite numbers. These ordinals give us a way to track the hydra's structure so that every move actually decreases its value, even when it seems to grow. This is what ultimately proves why the hydra sequnce ends.

However, the story does not end there. The proof relies on reasoning that goes beyond what first order Arithmetic can justify. This means that, although the statement is true, it cannot be proved within that system. The hydra game therefore becomes more than just a puzzle—it is a concrete example of incompleteness in mathematics.

In the end, the hydra teaches an unexpected lesson: even to understand why a finite process must eventually stop, we sometimes need to use ideas that reach into the infinite.

References

[1] [1] L. Kirby and J. Paris, "Accessible independence results for Peano arithmetic," *Bull. Lond. Math. Soc.* 14 1982. , 1.2)

[2] [2] L. Carlucci, "A new proof-theoretic proof of the independence of Kirby–Paris' hydra theorem," *Theor. Comput. Sci.* 300 2003.)

[3] [3] H. Tigerschiöld, "Introduction and proof of the Goodstein sequence and hydra game," 2017. , 3.5)

[4] [4] M. Hamano and M. Okada, "A relationship among Gentzen's proof-reduction, Kirby–Paris hydra game, and Buchholz's hydra game," 1995.)