BRIDGES
Mathematical Connections
in Art, Music, and Science

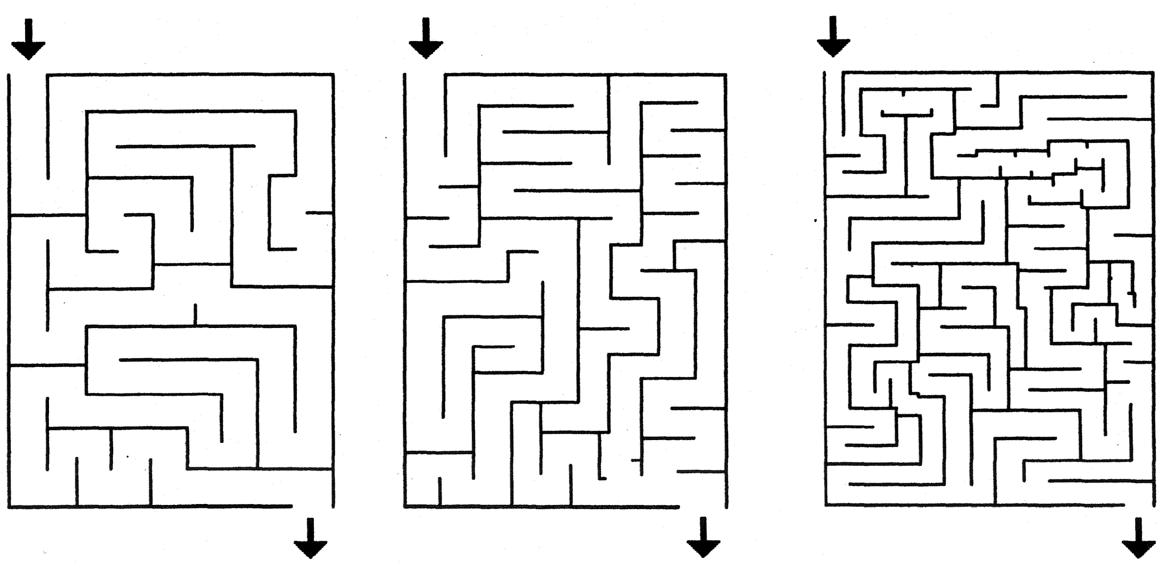
The Complexity and Difficulty of a Maze

Michael Scott McClendon
University of Central Oklahoma
Department of Mathematics and Statistics
100 North University Drive
Edmond, Oklahoma 73034-5209, U.S.A.
Email: mmcclendon@ucok.edu

Abstract

Some mazes are more difficult to solve than other mazes. For this reason, we develop a method by which the difficulty of a maze can be quantified. In the process we determine a way in which the complexity of the hallways in a maze, which is the degree to which a maze has short and quick twists and turns, can also be measured. We use the various complexity measures of the hallways in a maze in order to calculate the overall complexity and difficulty of the maze. We provide several examples in order to help establish some validity of the formulas developed in this paper. As the main tool used in developing our methods is continuum theory, we will use appropriate definitions throughout this paper.

1. Introduction



a. b. c.

Figure 1: *Three mazes of varying difficulty.*

Consider the mazes in figure 1. Of these three mazes, which one is the more challenging? If we can order these mazes according to this criterion, then the possibility of such an ordering of all mazes seems plausible. Given two mazes, we say the more "challenging" of the two has a higher measure of difficulty. In this paper we will develop a method by which the difficulty of a maze can be measured, and then provide some interesting examples. For this paper we will assume that all mazes can be embedded in the plane.

2. The Graph of a Maze

A *graph* is a continuum that can be written as the union of finitely many arcs any two of which are either disjoint or intersect only in one or both of their endpoints [2]. Let M be a maze. We want to define a function that maps the maze M onto a graph. Every hallway of the maze has two walls with a pathway between them. Let a be any point on one wall of a hallway and let b be the point on the other wall nearest to a. The line segment connecting a and b is called a *pathway perpendicular* of the maze. A *graph of a maze* is any graph that lies in the interior of the maze that satisfies the following:

1) The graph intersects each pathway perpendicular of the maze in the middle half of the pathway perpendicular.
2) No pathway perpendicular of the maze intersects the graph in more than one point.

Note that more than one graph can be derived from any given maze. Unless otherwise specified, we will refer to a graph of a maze simply as a maze. For example, we derive in figure 2 the graph of the maze from figure 1b.

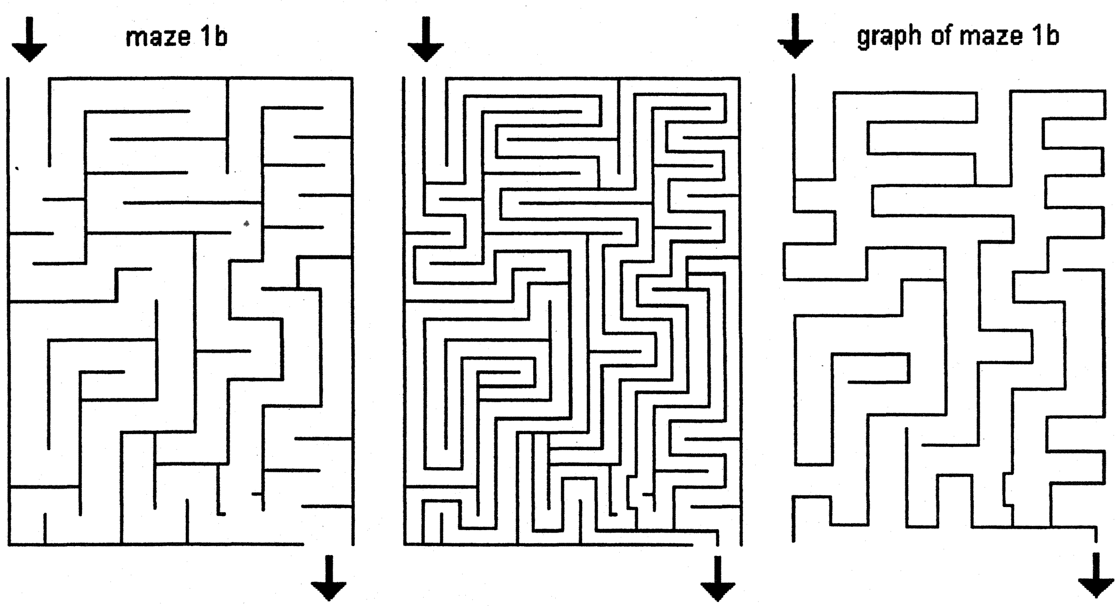


Figure 2: *Constructing the graph of a maze.*

We will also assume that every maze contains two points, called the *gates*, which are distinguished from all other points in the graph. Let p and q be the gates of a maze M. Though it may not be necessary, it is customary to consider p and q as an ordered pair (p,q). In this case we call the gate p the *entrance* or the *start* of M and we call the gate q the *exit* or the *finish* of M.

We denote the order of a point x in a maze by *ord*(x). If e is a point in a maze M such that ord(e) = 1 then e is called a *dead-end* of M. Now consider the subset K of M defined by $K = \{p \in M \mid \operatorname{ord}(p) \leq 2\}$. The components of K are called *hallways*. Every point in $M - K$ is called an *intersection*. We say two hallways h_i and h_j are *adjacent* if and only if there is a point $x \in \mathrm{M} - K$ such that $h_i \cup \{x\} \cup h_j$ is a connected set. If u and v are intersections or dead-ends such that $\{u\} \cup h \cup \{v\}$ is a connected set then we call u and v the *endpoints* of h. In this case, the set $\{u\} \cup h \cup \{v\}$ is called a *closed hallway*, and is denoted $\bar{h}$. A sequence of closed hallways $\bar{h}_1, \bar{h}_2, \ldots, \bar{h}_n$ is called a *trail* if and only if h_i is adjacent to h_{i+1} whenever $1 \le i \le n-1$. A *loop* is defined to be any trail that satisfies either of the following two conditions:

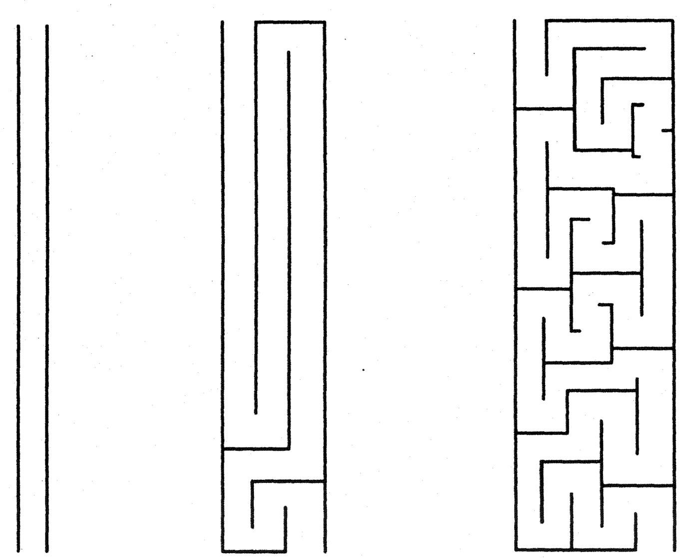
1) The trail consists of one hallway and the two endpoints of the hallway are the same point.
2) The trail consists of more than one hallway and each endpoint of each hallway in the trail is also an endpoint of exactly one other hallway in the trail.

We will assume throughout the rest of this paper that no maze contains a loop.

A *simple trail* is a trail in which no hallway is repeated. A *reduced trail* is a simple trail in which no intersection is repeated. Since we are not allowing mazes to have loops, then a trail is simple if and only if it is reduced (try proving this). A solution of a maze is any reduced trail whose endpoints are the gates of the maze. A maze is said to be *well-constructed* if it has exactly one solution. Again, since we are not allowing loops, then throughout the rest of this paper all mazes will be well-constructed. Note that solving a maze requires selecting the correct direction to travel at each intersection. However, if a maze has no loops, as the mazes that we are considering, then taking the "immediate left" direction at each intersection and reversing direction at each dead-end will always provide a trail (not necessarily a simple trail) from the start to the finish. Clearly, this trail will contain the solution of the maze.

Let M be a maze and suppose that the trail $T \subset M$ is the solution of the maze. Let $I = \{v_1, v_2, \ldots, v_n\}$ be the set of intersections located on T. Then any component of $M - T$ is called a *branch* of the maze. Each branch in M is connected to T by a point in I.

3. The Complexity Measure of a Hallway



a. b. c.

Figure 3: *Three hallways of varying complexity.*

Consider the three hallways in figure 3. As it is possible to order these hallways by how "complicated" or "confusing" each of these hallways are we want to develop a method by which we could order all hallways accordingly. Given two mazes, we say the more "complicated" of the two has a higher measure of complexity. It appears that the more quickly a hallway alters its direction the higher the measure of its complexity, or the more complex the hallway becomes. Therefore, the measure of the complexity of a hallway will depend on the magnitude of its direction changes and on how "quickly" it changes direction.

Let h be a hallway with endpoints u and v. Since h is connected and every point in h is of order 2 then h is homeomorphic to an arc. So let h have arclength $D(h)$ and let $f_h : [0,1] \to \mathbb{R}^2$ be a continuous function from the unit interval onto h such that $f_h(0) = u$, $f_h(1) = v$ and such that $f_h(t)$ is the point on h where the arclength of the subarc of h with endpoints u and $f_h(t)$ has length $tD(h)$. Note that $df_h / dt \neq 0$ on $[0,1]$. Now, the derivative df_h / dt evaluated at $t = t_0$ is the velocity vector of f_h at $t = t_0$. If we view the hallway h as the path of a particle moving from u to v, then the velocity vector at $t = t_0$ points in the direction that the particle is moving at time t_0. For all $t \in [0,1]$ define $V_+(t)$ to equal the one-sided derivative of f_h from the right of t and $V_-(t)$ to equal the one-sided derivative of f_h from the left of t. For all values of $t \in (0,1)$ for which it exists, define $V(t) = df_h / dt = V_+(t) = V_-(t)$.

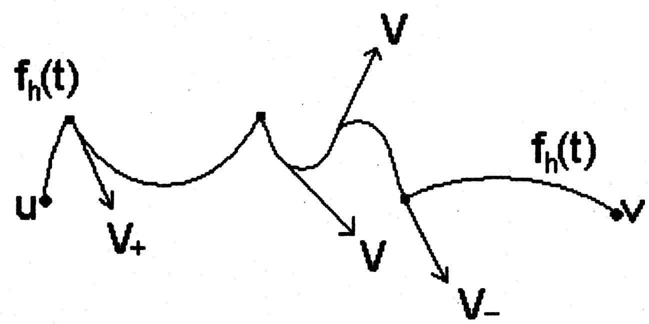


Figure 4: *A hallway and various direction vectors* $V_{\pm}(t)$

We will construct a subset W_h of points belonging to h. If it exists, let $t_1 \in (0,1]$ be the smallest value of t such that the direction of $V_{\pm}(t_1)$ differs from the direction of $V_+(0)$ by at least $\pi/4$. Let $w_1 = f_h(t_1) \in W_h$. We call the direction of $V_+(0)$ the *previous direction relative to* w_1. Define $\beta(w_1)$ to be equal to the absolute value of the difference in the radian measures between the directions of $V_-(t_1)$ and $V_+(t_1)$. That is, $\beta(w_1) = \left|\tan^{-1}(V_-(t_1)) - \tan^{-1}(V_+(t_2))\right|$. If $\beta(w_1) < \frac{\pi}{4}$, then we say that the *current direction at* w_1 is the direction of $V_-(t_1)$. Otherwise, we say that the current direction at w_1 is the direction of $V_+(t_1)$. If it exists, let $t_2 \in (t_1, 1]$ be the smallest value of t such that the direction of $V_{\pm}(t_2)$ differs from the current direction at w_1 by at least $\pi/4$. Let $w_2 = f_h(t_2) \in W_h$. We call the current direction at w_1 the *previous direction relative to* w_2. Define $\beta(w_2)$ to be equal to the absolute value of the difference in the radian measures between the directions of $V_-(t_2)$ and $V_+(t_2)$. If $\beta(w_2) < \frac{\pi}{4}$, then we say that the *current direction* at w_2 is the direction of $V_-(t_2)$. Otherwise, we say that the current direction at w_2 is the direction of $V_+(t_2)$. Inductively, once $w_1, w_2, \ldots, w_n$ have been selected, let $w_{n+1} = f_h(t_{n+1})$ be the point on h between w_n and v where t_{n+1} is the smallest value in $(t_n, 1]$ such that the direction of $V_{\pm}(t_{n+1})$ differs from the current direction at w_n by at least $\pi/4$.

Define $\theta(w_i)$ to be equal to the absolute value of the difference in the radian measures between the directions of $V(t_i)$ and $V(t_{i+1})$. That is, $\beta(w_i) = \left|\tan^{-1}(V(t_i)) - \tan^{-1}(V(t_{i+1}))\right|$. Roughly, θ is a function $\theta : W_h \to (0, \pi)$ that measures the overall change in direction of h as h crosses each point in

W_h. Denote the arc length of *h* between w_{i-1} and w_i in W_h by $d(w_i)$. We denote the arc length of *h* by *D*(*h*).

Suppose *h* is a hallway and that $W_h = \{w_1, w_2, \ldots, w_n\}$. The more extreme each turn is at each point w_i, the more complex *h* is. That is, the larger the value of each $\theta(w_i)$, the more complex *h* is. Therefore, the complexity of *h* increases as each value of $\theta(w_i)$ increases. Also, the shorter each subset of *h* is between points in W_h, the more complex *h* is. That is, the smaller the value of each $d(w_i)$, the more complex *h* is. Therefore, the complexity of *h* increases as each value of $\frac{1}{d(w_i)}$ increases. What makes a hallway most complex is when both $\theta(w_i)$ and $\frac{1}{d(w_i)}$ are simultaneously large. Finally, the longer the entire hallway, the more complex it is, so that the complexity of *h* increases as *D*(*h*) increases. For these reasons, we define the *complexity* $\gamma(h)$ of *h* from endpoint *u* to endpoint *v* by

$$\gamma(h) = D(h)\sum_{i=1}^{n}\frac{\theta(w_i)}{d(w_i)\cdot\pi}. \tag{1}$$

We divide by π to simplify the value. Also note that since the units of distance appear in both the numerator and the denominator, then $\gamma(h)$ is independent of the unit measurement of length. The hallways in figures 3a, 3b and 3c have complexities of 0, 138 and 2432, respectively.

4. The Complexity and Difficulty Measures of a Maze

Let *M* be a maze. Let *T* be the solution of the maze and suppose $B = \{h_1, h_2, \ldots, h_n\}$ is the set of all hallways in some branch of *M*. We define the complexity $\gamma(B)$ of *B* by

$$\gamma(B) = \sum_{i=1}^{n}\gamma(h_i) \tag{2}$$

where $\gamma(h_i)$ is the complexity of h_i from *u* to *v* where *u* lies between *T* and *v*.

Suppose $K = \{B_1, B_2, \ldots, B_n\}$ is the set of branches in a maze *M* with solution *T*. We define the *complexity* $\gamma(M)$ of *M* by

$$\gamma(M) = \log\left[\gamma(T) + \sum_{i=1}^{n}\gamma(B_i)\right] \tag{3}$$

where $\gamma(T)$ is the complexity of $T \subset M$ from start to finish and $\gamma(B_i)$ is the complexity of branch B_i. The mazes in figures 1a, 1b and 1c have complexities of 3.3, 3.1 and 3.4, respectively.

Note that the complexity measure of a maze *M* is an intrinsic measure. That is, in order to calculate $\gamma(M)$, we need to know the solution *T* so that we can calculate the value of $\gamma(T)$. It is possible to approximate $\gamma(M)$ very well by an extrinsic form of (3) that does not require knowing *T*. Suppose that $B = \{h_1, h_2, \ldots, h_n\}$ is the set of all hallways in a maze *M*. Then

$$\gamma(M) \approx \log\left[\sum_{i=1}^{n}\gamma(h_i)\right]$$

where $\gamma(h_i)$ is the complexity of h_i from one endpoint to the other endpoint.

Let the maze *M* have a solution *T*. If *T* has a small complexity, but also has many branches of large complexity, then $\gamma(M)$ would be large, but the maze would be easy to solve. If instead of adding the terms in (3) above, we were to multiply them, then we would arrive at a measure that seems to better describe the difficulty of a maze than the complexity measure. We add 1 to the complexity measure of each branch so that a single branch of some complexity less than one will not lower the overall complexity of a maze. That is, we define the difficulty measure $\delta(\mathrm{M})$ of M by

$$\delta(M) = \log\left[\gamma(T)\prod_{i=1}^{n}\left(\gamma(B_i)+1\right)\right]. \tag{4}$$

The mazes in figures 1a, 1b and 1c have difficulty measures of 3.3, 12.5 and 14.7, respectively.

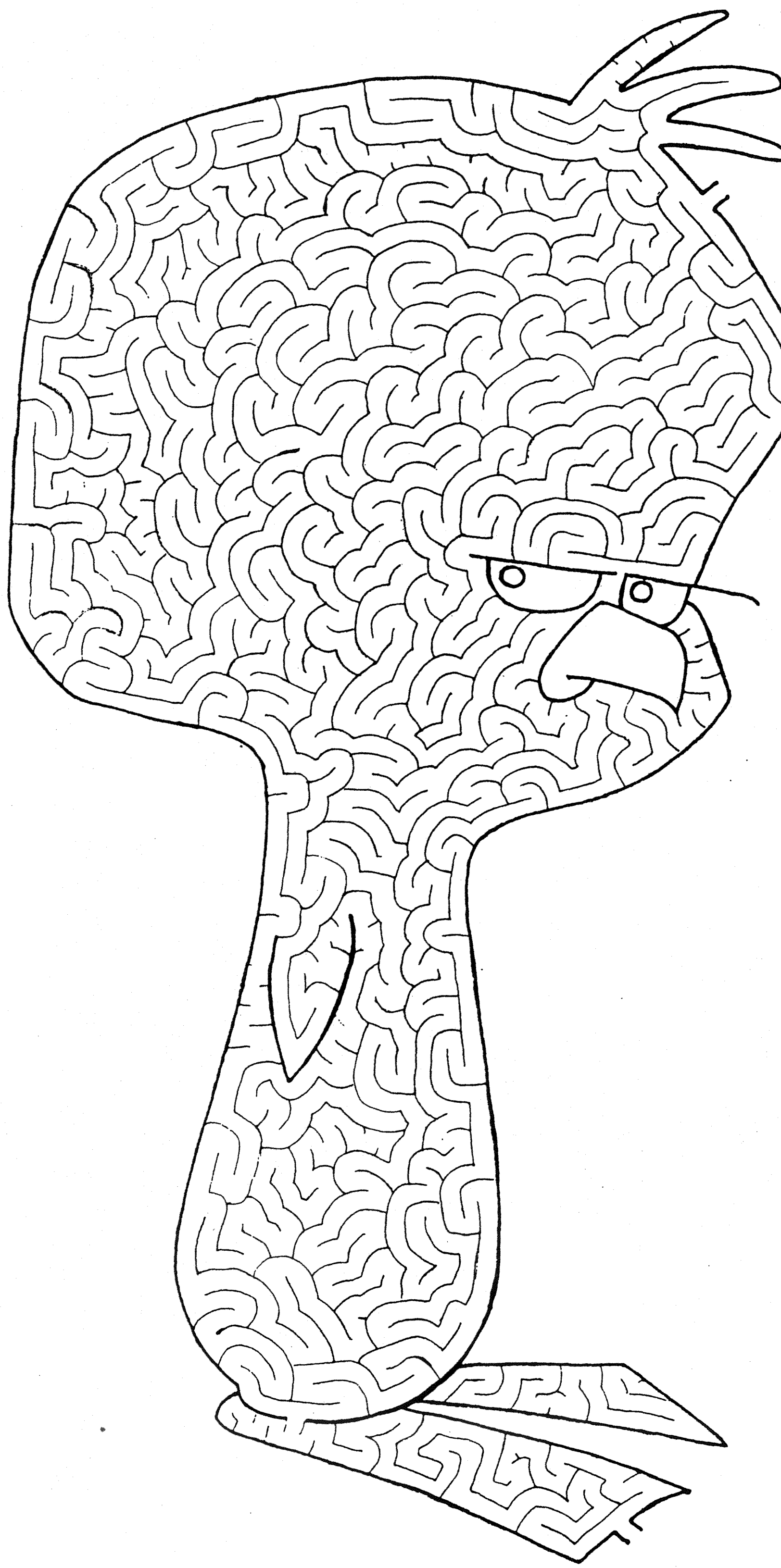
Note that we have defined the complexity and the difficulty of a maze while using the term "maze" to actually mean "graph of a maze". Recall that an actual maze can be associated with more than one graph. Given an actual maze M, suppose that G is the collection of all graphs that can be associated with M. Then we define $\gamma(M)$ and $\delta(M)$ to equal the infimum of $\gamma(g)$ and the infimum of $\delta(g)$, respectively, as g varies over every graph in G. That is, $\gamma(M) = \inf_{g \in G} \gamma(g)$ and $\delta(M) = \inf_{g \in G} \delta(g)$. Thus, we may not be able to practically calculate the complexity and difficulty of a maze, but we can calculate a reasonable upper bound for these measures.

Suppose that the points in W_h are selected so that the current direction at each $w_i \in W_h$ differs from the previous direction relative to w_i by say $\frac{\pi}{8}$ instead of $\frac{\pi}{4}$. It is an open question of determining precisely how the complexity measure of a maze will be altered. Another open problem would be to determine whether it is possible to construct a maze-generation algorithm that, given values γ and δ, yields a maze with complexity γ and difficulty δ. It is further unknown whether there is a maze that can be drawn in a finite space that has an infinite measure of complexity or an infinite measure of difficulty. Finally, for a maze M, note that $\delta(M)$ is an intrinsic measure, as is $\gamma(M)$. We are unsure if there is any way to approximate $\delta(M)$ extrinsically, as we can $\gamma(M)$.

Now we will provide some examples of mazes along with upper bounds of their associated measures of complexity and measures of difficulty. It should be evident that these measures do a pretty good job of identifying the complexity of a maze as well as predicting the difficulty of the maze. It should also be fun testing the validity of the measure of difficulty as defined in this paper.

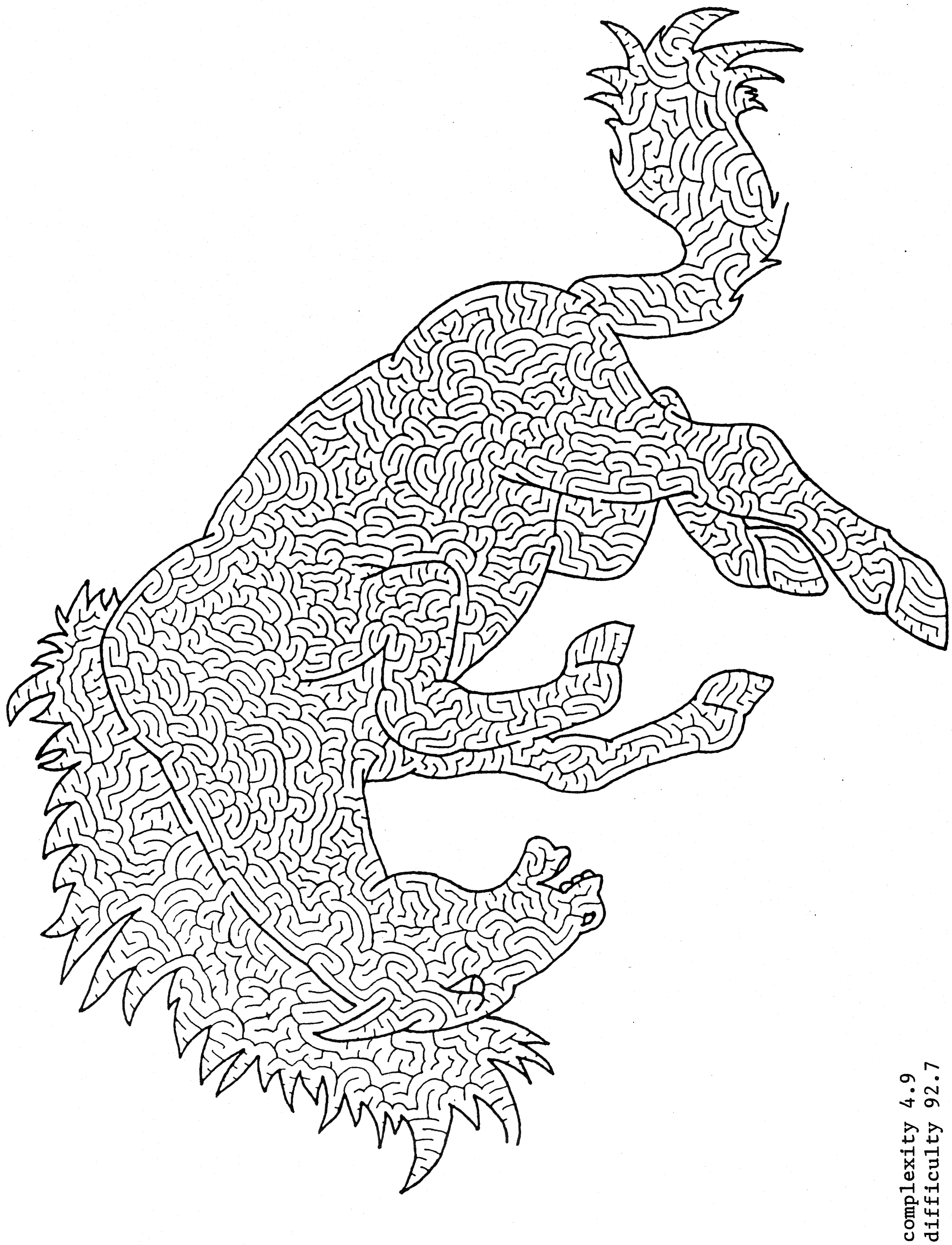
References

[1] J. Munkres, *Topology, A First Course*, Prentice Hall, New Jersey, 1975.
[2] S. Nadler, *Continuum Theory, An Introduction*, Marcel Dekker, New York, 1992.
[3] P. Anthony, "The Topology of Roman Mosaic Mazes", in *The Visual Mind: Art and Mathematics*, edited by M. Emmer, MIT Press, 1994, pp. 65 – 73.

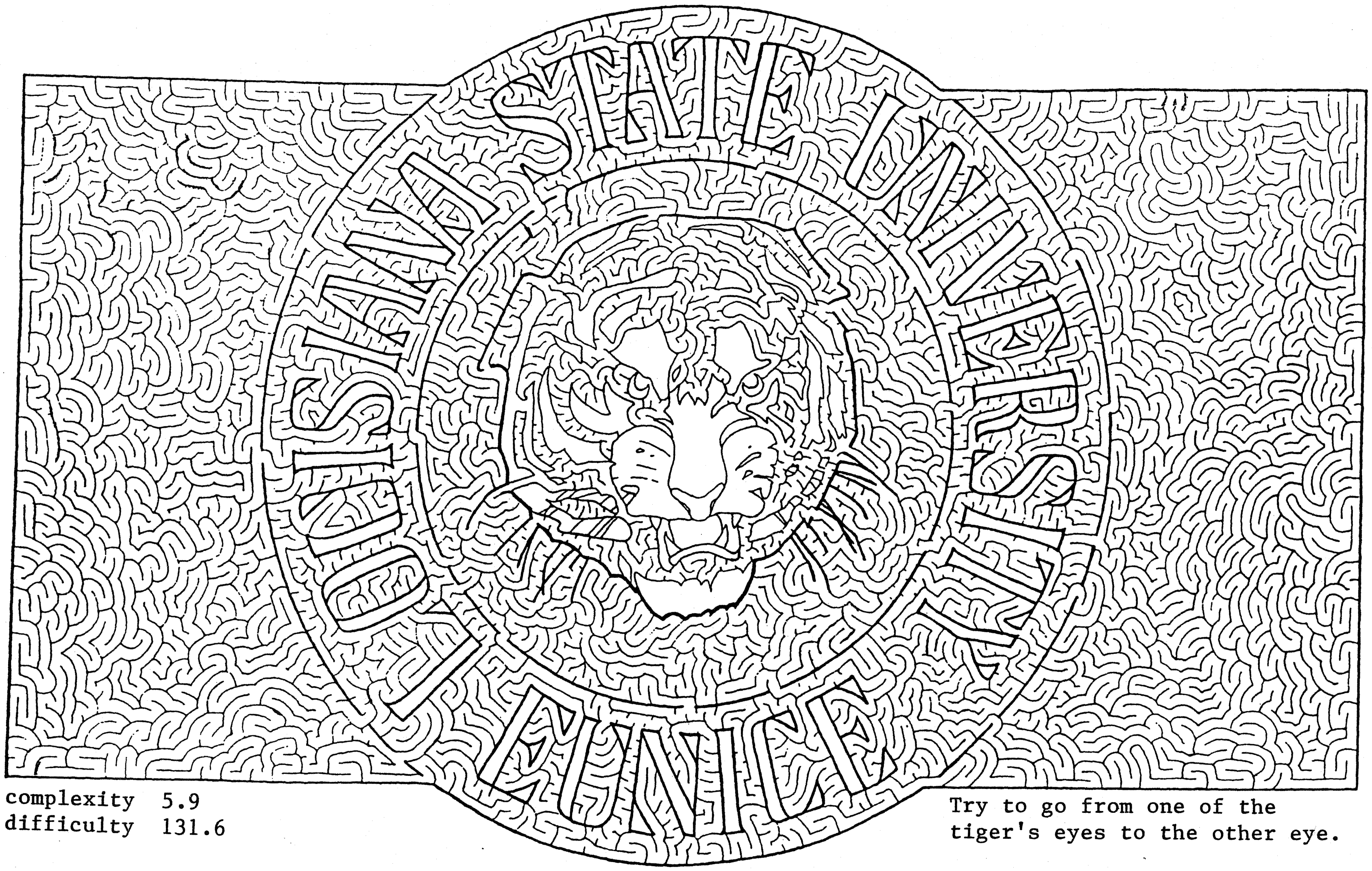


complexity 4.3
difficulty 55.4

complexity 4.8
difficulty 69.2



complexity 4.9
difficulty 92.7



complexity 5.9
difficulty 131.6

Try to go from one of the tiger's eyes to the other eye.