Asymptotics of minimal deterministic finite automata recognizing a finite binary language

AofA 2020

Andrew Elvey Price, Wenjie Fang, and Michael Wallner

Institut Denis-Poisson, Université de Tours, France

September, 2020
What is a DFA?
Deterministic finite automata (DFA)

DFA on alphabet \( \{a, b\} \)

Graph with

- two outgoing edges from each node (state), labelled \( a \) and \( b \)
- An initial state \( q_0 \)
- A set \( F \) of final states (coloured green).

Figure: A DFA.
Deterministic finite automata (DFA)

DFA on alphabet \{a, b\}

Graph with
- two outgoing edges from each node (state), labelled \(a\) and \(b\)
- An initial state \(q_0\)
- A set \(F\) of final states (coloured green).

Properties
- **Language:** the set of accepted words
- **Minimal:** no DFA with fewer states accepts the same language
- **Acyclic:** no cycles (except loops at unique sink)

Figure: A DFA. This is the minimal DFA recognising the language \{aa, aab, ab, b, bb\}. 

Elvey Price, Fang, Wallner | Bordeaux, Paris | 29.9.2020
Counting minimal acyclic DFAs

**This work:** Asymptotics of the numbers $m_n$ of minimal, acyclic DFAs on a binary alphabet with $n + 1$ nodes.

- Studied by Domaratzki, Kisman, Shallit and Liskovets between 2002 and 2006
- Best bounds were out by an exponential factor
- We gave upper and lower bounds differing by a $\Theta(n^{1/4})$ factor, by relating the DFAs to compacted trees.
Main result
Main result – A stretched exponential appears

Theorem

The number $m_n$ of minimal DFAs recognising a finite binary for $n \to \infty$

$$m_n = \Theta \left( n! 8^n e^{3a_1 n^{1/3}} n^{7/8} \right),$$

where $a_1 \approx -2.3381$ is the largest root of the Airy function

$$\text{Ai}(x) = \frac{1}{\pi} \int_0^\infty \cos \left( \frac{t^3}{3} + xt \right) dt.$$
Main result – A stretched exponential appears

Theorem

The number $m_n$ of minimal DFAs recognising a finite binary for $n \to \infty$

$$m_n = \Theta \left( n! \ 8^n e^{3a_1 n^{1/3}} n^{7/8} \right),$$

where $a_1 \approx -2.3381$ is the largest root of the Airy function

$$\text{Ai}(x) = \frac{1}{\pi} \int_0^\infty \cos \left( \frac{t^3}{3} + xt \right) dt.$$

Conjecture

Experimentally we find

$$m_n \sim \gamma n! 8^n e^{3a_1 n^{1/3}} n^{7/8},$$

where

$$\gamma \approx 76.438160702.$$
What is the Airy function?

**Properties**

- \( \text{Ai}(x) = \frac{1}{\pi} \int_0^\infty \cos \left( \frac{t^3}{3} + xt \right) dt \)
- Largest root \( a_1 \approx -2.3381 \)
- \( \lim_{x \to \infty} \text{Ai}(x) = 0 \)
- Also defined by \( \text{Ai}''(x) = x \text{Ai}(x) \)

- [Banderier, Flajolet, Schaeffer, Soria 2001]: Random Maps
- [Flajolet, Louchard 2001]: Brownian excursion area
Bijection to decorated paths
Bijection to decorated paths
Bijection to decorated paths

- Highlight spanning tree given by depth first search (ignoring the sink)
- I.e., Black path to each vertex is first in lexicographic order
- Colour other edges red
- Draw as a binary tree with $a$ edges pointing left and $b$ edges pointing right
Bijection to decorated paths

- Highlight spanning tree given by depth first search (ignoring the sink)
- I.e., Black path to each vertex is first in lexicographic order
- Colour other edges red
- Draw as a binary tree with $a$ edges pointing left and $b$ edges pointing right
Bijection to decorated paths

- Highlight spanning tree given by depth first search (ignoring the sink)
- I.e., Black path to each vertex is first in lexicographic order
- Colour other edges red
- Draw as a binary tree with $a$ edges pointing left and $b$ edges pointing right
Label nodes in post-order. By construction red edges point from a larger number to a smaller number

→ Label pointers
Label nodes in post-order. By construction red edges point from a larger number to a smaller number

→ Label pointers
Bijection to decorated paths

Elvey Price, Fang, Wallner | Bordeaux, Paris | 29.9.2020
Compacted Binary Trees | Bijection to decorated paths

Bijection to decorated paths

When the tree traversal...
- goes up: add up step with colour matching the corresponding node.
- passes a pointer:
  - add horizontal step
  - mark box corresponding to pointer label
Bijection to decorated paths

When the tree traversal...
- goes up: add up step with colour matching the corresponding node.
- passes a pointer:
  - add horizontal step
  - mark box corresponding to pointer label
Bijection to decorated paths

When the tree traversal...

- goes up: add up step with colour matching the corresponding node.
- passes a pointer:
  - add horizontal step
  - mark box corresponding to pointer label
When the tree traversal...

- goes up: add up step with colour matching the corresponding node.
- passes a pointer:
  - add horizontal step
  - mark box corresponding to pointer label
When the tree traversal...
- goes up: add up step with colour matching the corresponding node.
- passes a pointer:
  - add horizontal step
  - mark box corresponding to pointer label
Bijection to decorated paths

When the tree traversal...

- goes up: add up step with colour matching the corresponding node.
- passes a pointer:
  - add horizontal step
  - mark box corresponding to pointer label
Bijection to decorated paths

When the tree traversal...

- goes up: add up step with colour matching the corresponding node.
- passes a pointer:
  - add horizontal step
  - mark box corresponding to pointer label
When the tree traversal...

- goes up: add up step with colour matching the corresponding node.
- passes a pointer:
  - add horizontal step
  - mark box corresponding to pointer label
Bijection to decorated paths

When the tree traversal...
- goes up: add up step with colour matching the corresponding node.
- passes a pointer:
  - add horizontal step
  - mark box corresponding to pointer label
**Decorated paths**

- Path starts at \((-1, 0)\) and ends at \((n, n)\)
- Path stays below diagonal (after first step)
- One box is marked below each horizontal step
- Each vertical step is coloured white or green

By the bijection: The number of these paths is the number \(d_n\) of acyclic DFAs with \(n + 1\) nodes.
**Recurrence:** Denote by $a_{n,m}$ the number of paths ending at $(n, m)$.

\[ a_{n,m} = 2a_{n,m-1} + (m + 1)a_{n-1,m}, \quad \text{for } n \geq m \]

\[ a_{-1,0} = 1. \]

By the bijection: $d_n = a_{n,n}$ is the number of acyclic DFAs with $n + 1$ nodes.

What about minimality?
**Recurrence:** Denote by \( a_{n,m} \) the number of paths ending at \((n, m)\).  
\[
a_{n,m} = 2a_{n,m-1} + (m + 1)a_{n-1,m}, \quad \text{for } n \geq m
\]
\[
a_{-1,0} = 1.
\]

By the bijection: \( d_n = a_{n,n} \) is the number of acyclic DFAs with \( n + 1 \) nodes.

What about minimality?
For the DFA to be minimal, no state can be equivalent to a previous state:
- only possible if the new node is a leaf.
- If leaf is labelled $m + 1$, then $m$ choices of pointer labels and state colour must be avoided.
- Leaf corresponds to $\rightarrow \rightarrow \uparrow$ in path.
Minimal acyclic DFAs

For the DFA to be minimal, no state can be equivalent to a previous state:
- only possible if the new node is a leaf.
- If leaf is labelled $m + 1$, then $m$ choices of pointer labels and state colour must be avoided.
- Leaf corresponds to $\rightarrow \rightarrow \uparrow$ in path.
Minimal acyclic DFAs

For the DFA to be minimal, no state can be equivalent to a previous state:

- only possible if the new node is a leaf.
- If leaf is labelled $m+1$, then $m$ choices of pointer labels and state colour must be avoided.
- Leaf corresponds to $\rightarrow \rightarrow \uparrow$ in path.
Minimal acyclic DFAs

For the DFA to be minimal, no state can be equivalent to a previous state:
- only possible if the new node is a leaf.
- If leaf is labelled $m + 1$, then $m$ choices of pointer labels and state colour must be avoided.
- Leaf corresponds to $\rightarrow \rightarrow \uparrow$ in path.
Minimal acyclic DFAs

For the DFA to be minimal, no state can be equivalent to a previous state:
- only possible if the new node is a leaf.
- If leaf is labelled $m + 1$, then $m$ choices of pointer labels and state colour must be avoided.
- Leaf corresponds to $\rightarrow \rightarrow \uparrow$ in path.
Recurrence: Denote by $b_{n,m}$ the number of paths ending at $(n, m)$.

$$b_{n,m} = 2b_{n,m-1} + (m+1)b_{n-1,m} - mb_{n-2,m-1}, \quad \text{for } n \geq m,$$

$$b_{-1,0} = 1.$$  

By the bijection: $m_n = b_{n,n}$ is the number of minimal acyclic DFAs with $n + 1$ nodes.

Recurrence for minimal DFAs
Transforming recurrence for minimal DFAs

Recurrence: Denote by $b_{n,m}$ the number of paths ending at $(n, m)$.

\[ b_{n,m} = 2b_{n,m-1} + (m+1)b_{n-1,m} - mb_{n-2,m-1}, \quad \text{for } n \geq m, \]
\[ b_{-1,0} = 1. \]

By the bijection: $m_n = b_{n,n}$ is the number of minimal acyclic DFAs with $n+1$ nodes.

Transformation: Define $e_{n,m}$ by

\[ e_{n+m,n-m} = \frac{1}{n!2^{m-1}} b_{n,m}. \]

New recurrence:

\[ e_{n,m} = \frac{n-m+2}{n+m} e_{n-1,m-1} + e_{n-1,m+1} - \frac{n-m}{(n+m)(n+m-2)} e_{n-3,m-1}. \]

Now $m_n = n!2^{n-1}e_{2n,0}$.

Weights are now closer to 1, and steps (now ↗ and \downarrow) always increase $n$. 

Elvey Price, Fang, Wallner | Bordeaux, Paris | 29.9.2020
Transforming recurrence for minimal DFAs

Recurrence: Denote by $b_{n,m}$ the number of paths ending at $(n, m)$.

$$b_{n,m} = 2b_{n,m-1} + (m + 1)b_{n-1,m} - mb_{n-2,m-1}, \quad \text{for } n \geq m,$$

$$b_{-1,0} = 1.$$  

By the bijection: $m_n = b_{n,n}$ is the number of minimal acyclic DFAs with $n + 1$ nodes.

Transformation: Define $e_{n,m}$ by

$$e_{n+m,n-m} = \frac{1}{n!2^{m-1}} b_{n,m}.$$  

New recurrence:

$$e_{n,m} = \frac{n-m+2}{n+m} e_{n-1,m-1} + e_{n-1,m+1} - \frac{n-m}{(n+m)(n+m-2)} e_{n-3,m-1}.$$  

Now $m_n = n!2^{n-1}e_{2n,0}$.  

Weights are now closer to 1, and steps (now $\uparrow$ and $\downarrow$) always increase $n$.  

---

Elvey Price, Fang, Wallner | Bordeaux, Paris | 29.9.2020
Compacted Binary Trees | Bijection to decorated paths

Transforming recurrence for minimal DFAs

Recurrence: Denote by $b_{n,m}$ the number of paths ending at $(n, m)$.

$$b_{n,m} = 2b_{n,m-1} + (m + 1)b_{n-1,m} - mb_{n-2,m-1}, \quad \text{for } n \geq m,$$

$$b_{-1,0} = 1.$$

By the bijection: $m_n = b_{n,n}$ is the number of minimal acyclic DFAs with $n + 1$ nodes.

Transformation: Define $e_{n,m}$ by

$$e_{n+m,n-m} = \frac{1}{n!2^{m-1}} b_{n,m}.$$

New recurrence:

$$e_{n,m} = \frac{n-m+2}{n+m} e_{n-1,m-1} + e_{n-1,m+1} - \frac{n-m}{(n+m)(n+m-2)} e_{n-3,m-1}.$$

Now $m_n = n!2^{n-1}e_{2n,0}$.

Weights are now closer to 1, and steps (now ↗ and ↘) always increase $n$. 
Heuristics
Heuristics

We want to understand $e_{n,m}$ for large $n$.

Figure: Plots of $e_{n,m}$ against $m + 1$. **Left**: $n = 100$, **Right**: $n = 1000$

Let's zoom in to the left (small $m$) where interesting things are happening.

It seems to be converging to something...

**Guess**: $e_{n,m} \approx h(n)f\left(\frac{m + 1}{g(n)}\right)$. Moreover, we guess $g(n) = \sqrt[3]{n}$. 
Heuristics

We want to understand $e_{n,m}$ for large $n$.

![Plots of $e_{n,m}$ against $m + 1$.](image)

**Figure:** Plots of $e_{n,m}$ against $m + 1$. **Left:** $n = 100$, **Right:** $n = 1000$

- Let’s zoom in to the left (small $m$) where interesting things are happening.
- It seems to be converging to something...

**Guess:** $e_{n,m} \approx h(n)f\left(\frac{m + 1}{g(n)}\right)$. Moreover, we guess $g(n) = \sqrt[3]{n}$. 
Heuristics

We want to understand $e_{n,m}$ for large $n$.

![Plot of $e_{n,m}$ against $m + 1$. Left: $n = 100$, Right: $n = 1000$]

- Let's zoom in to the left (small $m$) where interesting things are happening.
- It seems to be converging to something...

**Guess:** $e_{n,m} \approx h(n)f\left(\frac{m + 1}{g(n)}\right)$. Moreover, we guess $g(n) = \sqrt[3]{n}$. 
We want to understand $e_{n,m}$ for large $n$.

**Figure:** Left: Plot of $e_{n,m}$ against $m+1$ for $n = 2000$. Right: Limiting function $f(x)$

- Let’s zoom in to the left (small $m$) where interesting things are happening.
- It seems to be converging to something...

**Guess:** $e_{n,m} \approx h(n)f\left(\frac{m + 1}{g(n)}\right)$. Moreover, we guess $g(n) = \sqrt[3]{n}$. 
Heuristics

We want to understand $e_{n,m}$ for large $n$.

Figure: **Left:** Plot of $e_{n,m}$ against $m + 1$ for $n = 2000$. **Right:** Limiting function $f(x)$

- Let’s zoom in to the left (small $m$) where interesting things are happening.
- It seems to be converging to something...

**Guess:** $e_{n,m} \approx h(n)f \left( \frac{m + 1}{g(n)} \right)$. Moreover, we guess $g(n) = \sqrt[3]{n}$. 
Heuristics

We want to understand $e_{n,m}$ for large $n$.

Figure: Left: Plot of $e_{n,m}$ against $m + 1$ for $n = 2000$. Right: Limiting function $f(x)$

- Let’s zoom in to the left (small $m$) where interesting things are happening.
- It seems to be converging to something...

Guess: $e_{n,m} \approx h(n)f\left(\frac{m + 1}{\sqrt[3]{n}}\right)$. Moreover, we guess $g(n) = \sqrt[3]{n}$. 
**Heuristic analysis of weighted paths**

**Recurrence:**

\[ e_{n,m} = \frac{n - m + 2}{n + m} e_{n-1,m-1} + e_{n-1,m+1} - \frac{n - m}{(n + m)(n + m - 2)} e_{n-3,m-1}. \]

**Guess:** \( e_{n,m} \approx h(n)f \left( \frac{m + 1}{\sqrt[3]{n}} \right) \).

Substitute into recurrence and set \( m = \kappa \sqrt[3]{n} - 1 \):

\[ s_n := \frac{h(n)}{h(n - 1)} \approx 2 + \frac{f''(\kappa) - 2\kappa f(\kappa)}{f(\kappa)} n^{-2/3} + O(n^{-1}) \]

Solution (assuming equality above):

\[ s_n = 2 + cn^{-2/3} + O(n^{-1}) \quad \Rightarrow \quad h(n) \approx 2^n e^{\frac{3c}{2} n^{1/3}} \]

\[ f''(\kappa) = (2\kappa + c)f(\kappa) \quad \Rightarrow \quad f(\kappa) = Ai(2^{-2/3}(2\kappa + c)) \]

Where \( c \) is constant.

Then \( f(0) = 0 \) implies \( c = 2^{2/3} a_1 \), where \( a_1 \approx -2.338 \) satisfies \( Ai(a_1) = 0 \).
Heuristic analysis of weighted paths

Recurrence:

\[ e_{n,m} = \frac{n - m + 2}{n + m} e_{n-1,m-1} + e_{n-1,m+1} - \frac{n - m}{(n + m)(n + m - 2)} e_{n-3,m-1}. \]

Guess: \( e_{n,m} \approx h(n) f \left( \frac{m + 1}{3\sqrt{n}} \right). \)

Substitute into recurrence and set \( m = \kappa \sqrt[3]{n} - 1: \)

\[ s_n := \frac{h(n)}{h(n-1)} \approx 2 + \frac{f''(\kappa) - 2\kappa f(\kappa)}{f(\kappa)} n^{-2/3} + O(n^{-1}) \]

Solution (assuming equality above):

\[ s_n = 2 + cn^{-2/3} + O(n^{-1}) \quad \Rightarrow \quad h(n) \approx 2^n e^{\frac{3\kappa}{2} n^{1/3}} \]

\[ f''(\kappa) = (2\kappa + c) f(\kappa) \quad \Rightarrow \quad f(\kappa) = Ai(2^{-2/3}(2\kappa + c)) \]

Where \( c \) is constant.

Then \( f(0) = 0 \) implies \( c = 2^{2/3} a_1 \), where \( a_1 \approx -2.338 \) satisfies \( Ai(a_1) = 0. \)
Heuristic analysis of weighted paths

Recurrence:

\[
e_{n,m} = \frac{n - m + 2}{n + m} e_{n-1,m-1} + e_{n-1,m+1} - \frac{n - m}{(n + m)(n + m - 2)} e_{n-3,m-1}.
\]

Guess:  \( e_{n,m} \approx h(n)f\left(\frac{m + 1}{\sqrt[3]{n}}\right). \)

Substitute into recurrence and set \( m = \kappa \sqrt[3]{n} - 1 \):

\[
s_n := \frac{h(n)}{h(n - 1)} \approx 2 + \frac{f''(\kappa) - 2\kappa f(\kappa)}{f(\kappa)} n^{-2/3} + O(n^{-1})
\]

Solution (assuming equality above):

\[
s_n = 2 + cn^{-2/3} + O(n^{-1}) \quad \Rightarrow \quad h(n) \approx 2^n e^{\frac{3c}{2} n^{1/3}}
\]

\[
f''(\kappa) = (2\kappa + c)f(\kappa) \quad \Rightarrow \quad f(\kappa) = Ai(2^{-2/3}(2\kappa + c))
\]

Where \( c \) is constant.

Then \( f(0) = 0 \) implies \( c = 2^{2/3}a_1 \), where \( a_1 \approx -2.338 \) satisfies \( Ai(a_1) = 0 \).
Heuristic analysis of weighted paths

Recurrence:

\[ e_{n,m} = \frac{n - m + 2}{n + m} e_{n-1,m-1} + e_{n-1,m+1} - \frac{n - m}{(n + m)(n + m - 2)} e_{n-3,m-1}. \]

Guess: \( e_{n,m} \approx h(n)f\left(\frac{m + 1}{\sqrt[3]{n}}\right) \).

Substitute into recurrence and set \( m = \kappa \sqrt[3]{n} - 1 \):

\[ s_n := \frac{h(n)}{h(n - 1)} \approx 2 + \frac{f''(\kappa) - 2\kappa f(\kappa)}{f(\kappa)} n^{-2/3} + O(n^{-1}) \]

Solution (assuming equality above):

\[ s_n = 2 + cn^{-2/3} + O(n^{-1}) \quad \Rightarrow \quad h(n) \approx 2^n e^{\frac{3c}{2} n^{1/3}} \]

\[ f''(\kappa) = (2\kappa + c)f(\kappa) \quad \Rightarrow \quad f(\kappa) = Ai(2^{-2/3}(2\kappa + c)) \]

Where \( c \) is constant and Ai is the Airy function.

Then \( f(0) = 0 \) implies \( c = 2^{2/3} a_1 \), where \( a_1 \approx -2.338 \) satisfies \( Ai(a_1) = 0 \).
Heuristic analysis of weighted paths

Recurrence:

\[ e_{n,m} = \frac{n - m + 2}{n + m} e_{n-1,m-1} + e_{n-1,m+1} - \frac{n - m}{(n + m)(n + m - 2)} e_{n-3,m-1}. \]

Guess: \( e_{n,m} \approx h(n)f \left( \frac{m + 1}{\sqrt[3]{n}} \right) \).

Substitute into recurrence and set \( m = \kappa \sqrt[3]{n} - 1 \):

\[ s_n := \frac{h(n)}{h(n - 1)} \approx 2 + \frac{f''(\kappa) - 2\kappa f(\kappa)}{f(\kappa)} n^{-2/3} + O(n^{-1}) \]

Solution (assuming equality above):

\[ s_n = 2 + cn^{-2/3} + O(n^{-1}) \quad \Rightarrow \quad h(n) \approx 2^n e^{\frac{3c}{2} n^{1/3}} \]

\[ f''(\kappa) = (2\kappa + c) f(\kappa) \quad \Rightarrow \quad f(\kappa) = \text{Ai}(2^{-2/3}(2\kappa + c)) \]

Where \( c \) is constant and \( \text{Ai} \) is the Airy function.

Then \( f(0) = 0 \) implies \( c = 2^{2/3} a_1 \), where \( a_1 \approx -2.338 \) satisfies \( \text{Ai}(a_1) = 0 \).
Refined heuristic analysis of weighted paths

Let \( a_1 \approx -2.3381 \) be the largest root of the Airy function \( \text{Ai} \).

**First guess:**

\[
e_{n,m} \approx h(n) f \left( \frac{m + 1}{\sqrt[3]{n}} \right),
\]

yields estimates

\[
h(n) \approx 2^n e^{3a_1(n/2)^{1/3}}
\]

\[
f(\kappa) = \text{Ai}(2^{1/3} \kappa + a_1)
\]

**Refined guess:**

\[
e_{n,m} \approx h(n) \left( f_0 \left( \frac{m + 1}{\sqrt[3]{n}} \right) + n^{-1/3} f_1 \left( \frac{m + 1}{\sqrt[3]{n}} \right) \right),
\]

yields estimates

\[
h(n) \sim \text{const} \cdot 2^n e^{3a_1(n/2)^{1/3}} n^{29/24}
\]

\[
f_0(\kappa) = \text{Ai}(2^{1/3} \kappa + a_1)
\]

This way we conjecture the asymptotic form for acyclic minimal DFAs:

\[
m_n = 2^{n-1} n! e_{2n,0} = \Theta \left( n! 8^n e^{3a_1 n^{1/3}} n^{7/8} \right)
\]
Refined heuristic analysis of weighted paths

Let $a_1 \approx -2.3381$ be the largest root of the Airy function $Ai$.

**First guess:**

$$e_{n,m} \approx h(n) f \left( \frac{m + 1}{\sqrt[3]{n}} \right),$$

yields estimates

$$h(n) \approx 2^n e^{3a_1(n/2)^{1/3}}$$

$$f(\kappa) = Ai(2^{1/3} \kappa + a_1)$$

**Refined guess:**

$$e_{n,m} \approx h(n) \left( f_0 \left( \frac{m + 1}{\sqrt[3]{n}} \right) + n^{-1/3} f_1 \left( \frac{m + 1}{\sqrt[3]{n}} \right) \right),$$

yields estimates

$$h(n) \sim \text{const} \cdot 2^n e^{3a_1(n/2)^{1/3}} n^{29/24}$$

$$f_0(\kappa) = Ai(2^{1/3} \kappa + a_1)$$

This way we conjecture the asymptotic form for acyclic minimal DFAs:

$$m_n = 2^{n-1} n! e_{2n,0} = \Theta \left( n! 8^n e^{3a_1 n^{1/3}} n^{7/8} \right)$$
Let $a_1 \approx -2.3381$ be the largest root of the Airy function $\text{Ai}$. 

**First guess:**

$$e_{n,m} \approx h(n)f\left(\frac{m+1}{\sqrt[3]{n}}\right),$$

yields estimates

$$h(n) \approx 2^n e^{3a_1(n/2)^{1/3}}$$

$$f(\kappa) = \text{Ai}(2^{1/3}\kappa + a_1)$$

**Refined guess:**

$$e_{n,m} \approx h(n)\left(f_0\left(\frac{m+1}{\sqrt[3]{n}}\right) + n^{-1/3}f_1\left(\frac{m+1}{\sqrt[3]{n}}\right)\right),$$

yields estimates

$$h(n) \sim \text{const} \cdot 2^n e^{3a_1(n/2)^{1/3}} n^{29/24}$$

$$f_0(\kappa) = \text{Ai}(2^{1/3}\kappa + a_1)$$

This way we conjecture the asymptotic form for acyclic minimal DFAs:

$$m_n = 2^{n-1} n! e_{2n,0} = \Theta\left(n! 8^n e^{3a_1 n^{1/3}} n^{7/8}\right)$$
Inductive proof
Proof method

Recall:

\[ e_{n,m} = \frac{n - m + 2}{n + m} e_{n-1,m-1} + e_{n-1,m+1} - \frac{n - m}{(n + m)(n + m - 2)} e_{n-3,m-1} \]

Number of minimal acyclic DFAs is \( m_n = 2^{n-1} n! e_{2n,0} \).

Method:

Find sequences \( A_{n,k} \) and \( B_{n,k} \) with the same asymptotic form, such that

\[ A_{n,k} \leq e_{n,k} \leq B_{n,k}, \]

for all \( k \) and all \( n \) large enough.
Proof method

Recall:
\[ e_{n,m} = \frac{n - m + 2}{n + m} e_{n-1,m-1} + e_{n-1,m+1} - \frac{n - m}{(n + m)(n + m - 2)} e_{n-3,m-1} \]

Number of minimal acyclic DFAs is \( m_n = 2^{n-1} n! e_{2n,0} \).

Method:
Find sequences \( A_{n,k} \) and \( B_{n,k} \) with the same asymptotic form, such that
\[ A_{n,k} \leq e_{n,k} \leq B_{n,k}, \]
for all \( k \) and all \( n \) large enough.

How to find them?

1. Use heuristics
2. Fiddle until they satisfy the recurrence of \( e_{n,k} \) with the equalities replaced by inequalities:
   \[ \quad \quad \quad \quad \quad \quad \quad = \quad \rightarrow \quad \leq \quad \text{and} \quad \geq \]
3. Prove \( A_{n,k} \leq e_{n,k} \leq B_{n,k} \) by induction.
Technicalities

Lots of technicalities:

- Before induction, we have to remove the negative term from the recurrence, but we have to do so precisely for asymptotics to stay the same.
- We only prove bounds for small $m$; we prove that large $m$ terms don’t matter.
- The lower bound is negative for very large $m$, so we have to be careful with induction.
- We only prove the bounds for sufficiently large $n$, but this only makes a difference to the constant term. Proof involves colourful Newton polygons:
Enumeration of minimal acyclic DFAs

1. Bijection to decorated paths
2. Recurrence for decorated paths
3. Heuristic analysis of recurrence
4. Inductive proof using heuristics

**Lower bound:**

\[ m_n \geq \gamma_1 n! 8^n e^{3a_1 n^{1/3}} n^{7/8}, \]

for some constant \( \gamma_1 > 0 \).
Summary

Enumeration of minimal acyclic DFAs

1. Bijection to decorated paths
2. Recurrence for decorated paths
3. Heuristic analysis of recurrence
4. Inductive proof using heuristics

Lower bound:

\[ m_n \geq \gamma_1 n!8^n e^{3a_1 n^{1/3}} n^{7/8}, \]

for some constant \( \gamma_1 > 0 \).

Upper bound (similar proof):

\[ m_n \leq \gamma_2 n!8^n e^{3a_1 n^{1/3}} n^{7/8}, \]

for some constant \( \gamma_2 > 0 \).
Theorem

The number of minimal DFAs recognizing a finite binary language satisfies for $n \to \infty$

$$m_n = \Theta \left( n! \, 8^n \, e^{3a_1} \frac{n^{1/3}}{8} \right),$$

where $a_1 \approx -2.3381$ is the largest root of the Airy function

$$Ai(x) = \frac{1}{\pi} \int_0^\infty \cos \left( \frac{t^3}{3} + xt \right) dt.$$ 

Further problems:

- Determining the constant term, or at least proving that one exists.
- How does the statistic number of states in DFA for a finite binary language interact with other natural statistics, like number of words? length of longest word? etc.
- For the method: Does anyone have a tricky recurrence to try?
Theorem

The number of minimal DFAs recognizing a finite binary language satisfies for $n \to \infty$

$$m_n = \Theta \left( n! \ 8^n \ e^{3a_1} n^{1/3} \ n^{7/8} \right),$$

where $a_1 \approx -2.3381$ is the largest root of the Airy function

$$\text{Ai}(x) = \frac{1}{\pi} \int_0^\infty \cos \left( \frac{t^3}{3} + xt \right) dt.$$ 

Further problems:

- Determining the constant term, or at least proving that one exists.
- How does the statistic number of states in DFA for a finite binary language interact with other natural statistics, like number of words? length of longest word? etc.
- For the method: Does anyone have a tricky recurrence to try?