Edge-Attachment Profiles of Graphs: An Exact Characterization

Brian Klemm

August 2026

Abstract

Classify each edge of a simple graph by how many of its endpoints it shares with another edge: *independent* (i) if neither, *partial* (p) if exactly one, *full* (f) if both. This collapses any graph on n edges to a triple (i, p, f) with $i + p + f = n$, which we call its **attachment profile**. We ask how many profiles are achievable.

We give a complete answer. Of the $\binom{n+2}{2}$ compositions of n into three non-negative parts, exactly five fail to be achieved for every $n \geq 3$, and we identify them explicitly. Hence

$$S(n) = \binom{n+2}{2} - 5, \qquad n \geq 3,$$

with $S(1) = 1$ and $S(2) = 2$. The proof is elementary and rests on a single observation: attachment is symmetric, so a graph cannot have exactly one attached edge.

We then define the **Edge Growth Automaton (EGA)**, a five-rule rewriting system on triples that adds one edge at a time without storing any adjacency structure, and prove it *sound and complete*: the states reachable in n steps are exactly the achievable profiles at n edges, with no omissions and no spurious states.

The resulting count coincides with OEIS A052905 under a shift of two; we do not claim the sequence as new. The contributions are the characterization, the soundness-and-completeness result, and a graph-theoretic interpretation of A052905 that does not appear in its entry.

1 Introduction

Graph enumeration usually means counting graphs, which means deciding isomorphism, which is expensive. A cheaper question is how much can be said about a graph from a coarse invariant that is trivial to compute.

The invariant here is local and about edges rather than vertices. Fix a simple graph G. For an edge $e = \{u, v\}$, define its **attachment number**

$$\alpha(e) = [\deg u \geq 2] + [\deg v \geq 2] \in \{0, 1, 2\},$$

counting how many of e's endpoints carry another edge. Sorting the edges of G by α gives the **attachment profile**

$$\pi(G) = (i, p, f), \qquad i + p + f = |E(G)|,$$

where i, p and f count the edges with $\alpha = 0, 1, 2$ respectively. Isolated vertices do not affect π, so we may assume G has none.

The profile is a drastic compression: it forgets which edges are attached to which, retaining only how many are attached how much. A natural question is what it retains. Write

$$\mathcal{P}(n) = \{\pi(G) : |E(G)| = n\}, \qquad S(n) = |\mathcal{P}(n)|.$$

Since $i + p + f = n$, we have $S(n) \le \binom{n+2}{2}$, the number of compositions of n into three non-negative parts. Section 3 shows the bound is tight up to a constant deficit of five.

Section 4 introduces an automaton that computes $\mathcal{P}(n)$ incrementally, one edge at a time, on triples alone — no adjacency lists, no isomorphism tests — and proves it computes exactly the right set.

1.1 Relationship to prior work

The count $S(n)$ agrees with OEIS A052905, $a(m) = (m^2 + 7m + 2)/2$, under $S(n) = a(n-2)$ for $n \ge 3$; A052905 has been catalogued since 2000. The sequence is therefore not new, and no novelty is claimed for it. What appears to be new is the characterization of *which* profiles are missing and why, the soundness-and-completeness result for the automaton, and the interpretation of A052905 as an edge-attachment count.

2 The Structure Theorem

Everything follows from grouping profiles by the number of **attached** edges,

$$d = p + f,$$

so that $i = n - d$. The results below say that $\mathcal{P}(n)$ is determined entirely by which (p, f) splits are possible at each d, independently of n.

Lemma 1 (Components)

The profile is additive over connected components. A component with one edge contributes $(1, 0, 0)$*; a component with* $k \ge 2$ *edges contributes* $(0, p, f)$ *with* $p + f = k$.

Proof. Additivity is immediate, since deg is computed within a component. For the second claim, suppose a component has $k \ge 2$ edges and contains an edge $e = \{u, v\}$ with $\alpha(e) = 0$. Then $\deg u = \deg v = 1$, so e meets no other edge and is a component by itself, contradicting $k \ge 2$. Hence every edge of such a component has $\alpha \ge 1$. ☐

So the independent edges are exactly the one-edge components, and d counts the edges lying in components of size ≥ 2.

Lemma 2 (Connected profiles)

Let C_k denote the set of pairs (p, f) realized by connected graphs with k edges. Then

$$C_2 = \{(2,0)\}, \qquad C_3 = \{(3,0),\,(2,1),\,(0,3)\}, \qquad C_k = \{(p,f) : p+f=k\} \text{ for } k \geq 4.$$

Proof. For $k = 2$ the only connected graph is the path on three vertices; its two edges each have one endpoint of degree 2 and one of degree 1, giving $(2, 0)$.

For $k = 3$ the connected graphs are the star $K_{1,3}$, the path P_4, and the triangle K_3. In the star every edge joins the degree-3 centre to a leaf, giving $(3, 0)$. In the path $a{-}b{-}c{-}d$, the outer edges have one endpoint of degree 1 and the middle edge has both endpoints of degree 2, giving $(2, 1)$. In the triangle every vertex has degree 2, giving $(0, 3)$. Note $(1, 2)$ is absent.

For $k \geq 4$ we construct each f from 0 to k, writing $p = k - f$:

f	construction	why it works
0	star $K_{1,k}$	every edge meets a leaf
1	edge uv, with $a \geq 1$ pendants at u and $b \geq 1$ at v, $a + b = k - 1$	only uv has both ends of degree ≥ 2; needs $k \geq 3$
2	path $u{-}v{-}w$, with $a \geq 1$ pendants at u and $b \geq 1$ at w, $a + b = k - 2$	uv and vw are full; needs $k \geq 4$
≥ 3	cycle C_f with $k - f$ pendant edges hung on its vertices	cycle edges are full, pendants partial

Each is connected and has exactly k edges, so all $k + 1$ splits occur. □

Theorem 1 (Characterization)

A triple (i, p, f) with $i + p + f = n$ lies in $\mathcal{P}(n)$ if and only if, writing $d = p + f$:

- $d = 0$*; or*
- $d = 2$ *and* $(p, f) = (2, 0)$*; or*
- $d = 3$ *and* $(p, f) \in \{(3, 0), (2, 1), (0, 3)\}$*; or*
- $d \geq 4$.

Equivalently, for $n \geq 3$ the profiles excluded are exactly the five

$$(n{-}1,\,1,\,0),\quad (n{-}1,\,0,\,1),\quad (n{-}2,\,1,\,1),\quad (n{-}2,\,0,\,2),\quad (n{-}3,\,1,\,2).$$

Proof. By Lemma 1, the attached edges decompose into components of size ≥ 2, so d must be a sum of parts each ≥ 2, and (p, f) the corresponding sum drawn from the C_k.

$d = 0$ is the all-independent profile $(n, 0, 0)$, realized by n disjoint edges.

$d = 1$ cannot be written as a sum of parts ≥ 2, so no profile with $p + f = 1$ occurs. This excludes $(n{-}1, 1, 0)$ and $(n{-}1, 0, 1)$.

$d = 2$ admits only the single part $k = 2$, and $C_2 = \{(2, 0)\}$. This excludes $(n{-}2, 1, 1)$ and $(n{-}2, 0, 2)$.

$d = 3$ admits only $k = 3$, since $3 = 2 + 1$ requires a forbidden part of size 1. So the splits are exactly C_3, excluding $(n{-}3, 1, 2)$.

$d \geq 4$: a single component of size d already realizes every split, by Lemma 2.

Padding with $n - d$ disjoint edges realizes each surviving case at n edges. $\square$

The reasoning behind the two hardest cases is worth stating without symbols, because it is the whole content of the theorem.

> **Attachment is symmetric.** A graph cannot contain exactly one attached edge — whatever that edge is attached to is itself attached. That kills $d = 1$.
>
> With exactly two attached edges, each is attached only to the other, hence at a single endpoint: both are partial. A full edge needs two distinct neighbours and cannot exist among only two attached edges. That kills every $d = 2$ split but $(2, 0)$.
>
> With exactly three, they form one connected three-edge graph — a star, a path, or a triangle — and those have profiles $(3, 0)$, $(2, 1)$, $(0, 3)$. Never $(1, 2)$.

Corollary 1 (Counting)

$$S(n) = \binom{n+2}{2} - 5 \quad (n \geq 3), \qquad S(1) = 1, \quad S(2) = 2.$$

Proof. There are $d{+}1$ triples with $i{+}p{+}f = n$ for each $d = 0, \dots, n$, totalling $\binom{n+2}{2}$. Theorem 1 removes 2 at $d = 1$, 2 at $d = 2$ and 1 at $d = 3$, all present once $n \geq 3$. For $n = 2$ only $d \leq 2$ occurs, removing 4 from 6; for $n = 1$ only $d \leq 1$, removing 2 from 3. $\square$

Corollary 2 (Closed forms)

For $n \geq 3$,

$$S(n) = \frac{n^2 + 3n - 8}{2} = T(n{+}1) - 5 = A052905(n-2), \qquad S(n) = S(n{-}1) + n + 1 \quad (n \geq 4),$$

where $T(k) = k(k + 1)/2$ are the triangular numbers (A000217). The first fifteen terms are

$$1,\ 2,\ 5,\ 10,\ 16,\ 23,\ 31,\ 40,\ 50,\ 61,\ 73,\ 86,\ 100,\ 115,\ 131.$$

The quadratic is therefore not an empirical fit but a consequence of a constant deficit: the number of compositions grows quadratically, and exactly five are lost at every $n \geq 3$.

3 Example: $n = 4$

The ten achievable profiles, with a witness graph for each.

profile	witness
$(4, 0, 0)$	four disjoint edges
$(2, 2, 0)$	a path on 3 vertices, plus two disjoint edges
$(1, 3, 0)$	star $K_{1,3}$, plus a disjoint edge
$(1, 2, 1)$	path P_4, plus a disjoint edge
$(1, 0, 3)$	triangle, plus a disjoint edge
$(0, 4, 0)$	star $K_{1,4}$
$(0, 3, 1)$	star $K_{1,3}$ with one edge subdivided
$(0, 2, 2)$	path P_5
$(0, 1, 3)$	triangle with one pendant edge (the "paw")
$(0, 0, 4)$	cycle C_4

The five excluded are $(3, 1, 0)$, $(3, 0, 1)$, $(2, 1, 1)$, $(2, 0, 2)$ and $(1, 1, 2)$, matching Theorem 1 with $n = 4$. Note $\binom{6}{2} - 5 = 15 - 5 = 10$.

4 The Edge Growth Automaton

Theorem 1 describes $\mathcal{P}(n)$ statically. We now give a generative description: a rewriting system on triples that builds $\mathcal{P}(n)$ from $\mathcal{P}(n-1)$ one edge at a time, touching no adjacency structure.

4.1 Definition

The **Edge Growth Automaton** starts from the single state $(1, 0, 0)$ and applies, at each step, any of the following five rules whose guard is satisfied. Each rule adds exactly one edge, so $i + p + f$ increases by one.

	rule	effect	guard	graph operation
R1	add independent	$(i{+}1,\ p,\ f)$	—	place a disjoint edge
R2	extend	$(i,\ p{+}1,\ f)$	$p > 0$ or $f > 0$	hang a new edge off an attached one

	rule	effect	guard	graph operation
R3	attach to independent	$(i{-}1,\ p{+}2,\ f)$	$i > 0$	join a new edge to a lone edge; both become partial
R4	promote	$(i,\ p,\ f{+}1)$	$p > 1$	extend a chain, filling its interior
R5	close	$(i,\ p{-}2,\ f{+}3)$	$p \geq 2$	close a cycle

Let R_n be the set of states reachable in $n-1$ steps from $(1, 0, 0)$.

4.2 Soundness and completeness

Theorem 2

$$R_n = \mathcal{P}(n) \quad \text{for all } n \geq 1.$$

The automaton reaches every achievable profile and no others.

Proof. Track $d = p + f$. The rules change d by $0, +1, +2, +1, +1$ respectively, so d **never decreases**, and R1 is the only rule leaving d fixed — and it leaves (p, f) untouched. Consequently the set of (p, f) splits available at each d can be analysed independently of i, and R1 supplies every larger i by padding.

No spurious states. The start state has $d = 0$. A state with $d = 0$ has $p = f = 0$, so R2, R4 and R5 are all blocked and only R1 ($\Delta d = 0$) and R3 ($\Delta d = +2$) apply. Hence $d = 1$ is never entered, and since d cannot decrease, no reachable state has $d = 1$. A state with $d = 2$ is reached either from $d = 0$, necessarily by R3 giving $(p, f) = (2, 0)$, or from another $d = 2$ state by R1, which preserves (p, f); so $(2, 0)$ is the only split at $d = 2$. From $(i, 2, 0)$ the rules yielding $d = 3$ are R2, R4 and R5, giving $(3, 0)$, $(2, 1)$ and $(0, 3)$ — exactly C_3, and in particular not $(1, 2)$. These are precisely the exclusions of Theorem 1.

No omissions. By induction on d, with R1 handling i. The bases $d \leq 3$ are the previous paragraph. Suppose every split at level $d \geq 3$ is reachable, and take a target (p', f') with $p' + f' = d + 1$.

- If $f' = 0$: apply R2 to $(d, 0)$.
- If $p' \geq 1$ and $(p' - 1, f')$ is a valid split at level d: apply R2 to it.
- The two exceptions at $d + 1 = 4$ are handled directly: $(2, 2)$ by R4 from $(2, 1)$, and $(0, 4)$ by R5 from $(2, 1)$.
- If $p' = 0$ and $f' = k \geq 5$: apply R5 to $(2, k - 3)$, a valid split at level $k - 1 \geq 4$.

For $d + 1 \geq 5$ the second case covers every split with $p' \geq 1$, since all splits at level $d \geq 4$ are valid, and the fourth covers $p' = 0$. $\square$

Machine verification: $R_n = \mathcal{P}(n)$ as sets — not merely $|R_n| = S(n)$ — for $n = 1, \ldots, 24$, where $\mathcal{P}(n)$ was computed independently by exhaustive graph enumeration for $n \leq 7$ and by component decomposition thereafter.

4.3 Minimality

Proposition 1

No proper subset of $\{R1, \dots, R5\}$ *generates* $\mathcal{P}(n)$ *for all* n.

Proof. For each rule we exhibit a profile that becomes unreachable without it.

- **R1.** No other rule increases i, and the start state has $i = 1$. So $(2, 0, 0)$ requires R1.
- **R3.** R3 is the only rule that decreases i. From $(1, 0, 0)$, the guards of R2, R4 and R5 all fail, so R1 and R3 are the only applicable rules; reaching $(0, 2, 0)$ requires R3.
- **R2.** Only R2 and R3 increase p, and R3 requires $i > 0$. From $(0, 2, 0)$, which has $i = 0$, R3 is blocked, so $(0, 3, 0)$ requires R2.
- **R4.** The reachable profiles at $d = 3$ are $(3, 0)$, $(2, 1)$ and $(0, 3)$, produced from $(i, 2, 0)$ by R2, R4 and R5 respectively. Neither R2 nor R5 yields $(2, 1)$, so $(0, 2, 1)$ requires R4.
- **R5.** R5 is the only rule that decreases p, and the only other rule increasing f is R4, whose guard $p > 1$ prevents it from ever reducing p to zero. So $(0, 0, 3)$ requires R5.

Each rule therefore contributes a profile no combination of the others produces. □

Every rule is load-bearing, and the five together are the smallest sound and complete system on triples.

5 A Cautionary Remark: When a Plausible Rule Outruns Its Objects

Earlier versions of this automaton carried a sixth rule,

$$(\text{convert } p \to f): \quad (i,\, p,\, f) \;\longmapsto\; (i,\, p{-}1,\, f{+}2), \qquad p > 0,$$

intended to model attaching a new edge to the free end of a partial edge, thereby "completing" it. The rule reads naturally and is easy to justify informally. It is nevertheless wrong, and it is worth exhibiting the failure because it is the kind of error the triple representation invites.

Take the path $a{-}b{-}c$ together with a disjoint edge: profile $(1, 2, 0)$ at $n = 3$. Now add an edge at c, producing $a{-}b{-}c{-}d$ plus the disjoint edge. In the real graph, ab and cd each have one endpoint of degree 1 and bc has both endpoints of degree 2, so the new profile is $(1, 2, 1)$. The rule instead reports $(1, 1, 2)$: it converts a partial edge to full and *invents a second full edge*, when in fact the edge just added is itself partial.

The damage is small and completely localized. Including the rule adds exactly one unachievable state at each $n \geq 4$, namely $(n{-}3, 1, 2)$ — the split ruled out by C_3 in Lemma 2 — and changes the count from $\binom{n+2}{2} - 5$ to $\binom{n+2}{2} - 4$, giving the sequence $1, 2, 5, 11, 17, 24, 32, \dots$ in place of $1, 2, 5, 10, 16, 23, 31, \dots$.

That the discrepancy is exactly one state per level, arising from exactly one rule, is what makes Theorem 2 worth stating. Soundness of an abstraction like this one is not automatic, and it is not visible from the growth rate: both rule sets grow quadratically with the same leading behaviour and differ only in a constant. Fitting a curve to either would have succeeded equally well.

6 Open Questions

1. **Finer profiles.** The profile discards a great deal — at $n = 4$ it merges the star $K_{1,4}$ with no other graph, but at larger n many non-isomorphic graphs share a profile. What is the size of the largest fibre $|\{G : \pi(G) = \tau\}|$, and which τ maximizes it? Refining α to record the degrees themselves, rather than only whether they exceed 1, gives a strictly finer invariant whose count is not known to us.
2. **Restricted classes.** Theorem 1 concerns all simple graphs. The analogous characterization for trees, for connected graphs, or for graphs of bounded degree should be tractable by the same component argument, and the deficits will differ.
3. **Multigraphs and hypergraphs.** The symmetry argument that kills $d = 1$ is robust and should survive both generalizations; the $d = 3$ analysis will not.
4. **Automata for other invariants.** Theorem 2 says a five-rule system on triples exactly captures a graph invariant. It is not clear how special this is. Which graph invariants admit a finite, sound and complete incremental rewriting system on their own values?

7 Conclusion

The attachment profile is about as coarse an invariant as a graph admits while still being defined edge-by-edge, and its image is almost everything: of the $\binom{n+2}{2}$ candidate triples, all but five occur, for every $n \geq 3$. The five that fail do so for reasons that are structural and easy to state — attachment is symmetric, a full edge needs two neighbours, and three mutually attached edges form a star, a path, or a triangle.

The resulting count is a known quadratic, A052905 under a shift; we claim the interpretation rather than the sequence. The Edge Growth Automaton computes the profile set incrementally on triples alone, and is sound and complete with respect to it, which is the sense in which the compression loses nothing it was meant to keep.

Appendix A — Verification

All claims were checked computationally in addition to being proved.

- Exhaustive enumeration of all **connected** graphs with $k \leq 7$ edges, over all vertex counts up to $k + 1$ (the bound for connectedness), computing π directly and confirming Lemma 2.
- Independent of Lemmas 1 and 2: exhaustive enumeration of **every** graph with $n \leq 5$ edges, over all vertex counts up to $2n$, confirming $\mathcal{P}(n)$ directly. The bound is $2n$ rather than $n + 1$ because n disjoint edges span $2n$ vertices.
- Explicit construction of the Lemma 2 witnesses for $k = 4, \dots, 30$, verified connected, of the right size, and of the right profile.
- Component-decomposition dynamic program computing $\mathcal{P}(n)$ to $n = 30$, checked against $\binom{n+2}{2} - 5$.
- Automaton reachability to $n = 201$, compared level by level against Theorem 1.
- Set-level equality $R_n = \mathcal{P}(n)$ for $n = 1, \dots, 24$.

- The witness profiles of Proposition 1, confirmed unreachable under each four-rule subsystem.
- $S(n) = A052905(n-2)$ verified algebraically and numerically for $n = 3, \dots, 40$.

Every result in this paper is proved; the computations above are corroboration and none of the claims depend on them. They are reproduced by a single self-contained script, `ega-verify.py`, which prints a pass/fail line for each bullet and requires no dependencies beyond Python 3.8.

Appendix B — The five excluded profiles

For reference, at $n \geq 3$:

profile	why impossible
(n−1, 1, 0)	exactly one attached edge; attachment is symmetric
(n−1, 0, 1)	same, and a full edge needs two neighbours
(n−2, 1, 1)	two attached edges are attached only to each other, so both are partial
(n−2, 0, 2)	same
(n−3, 1, 2)	three mutually attached edges form a star, path or triangle: $(3, 0)$, $(2, 1)$, $(0, 3)$