Syntax for clear Scientific exposition. 0. Preliminary Discussion

E. Anderson[1]

Abstract

This is an incipient discussion on a theory of Syntax for clear Scientific exposition. In which word count is replaced by concept count. And then by various simple-graph qualifiers in the sense of Graph Theory, at least to good approximation.

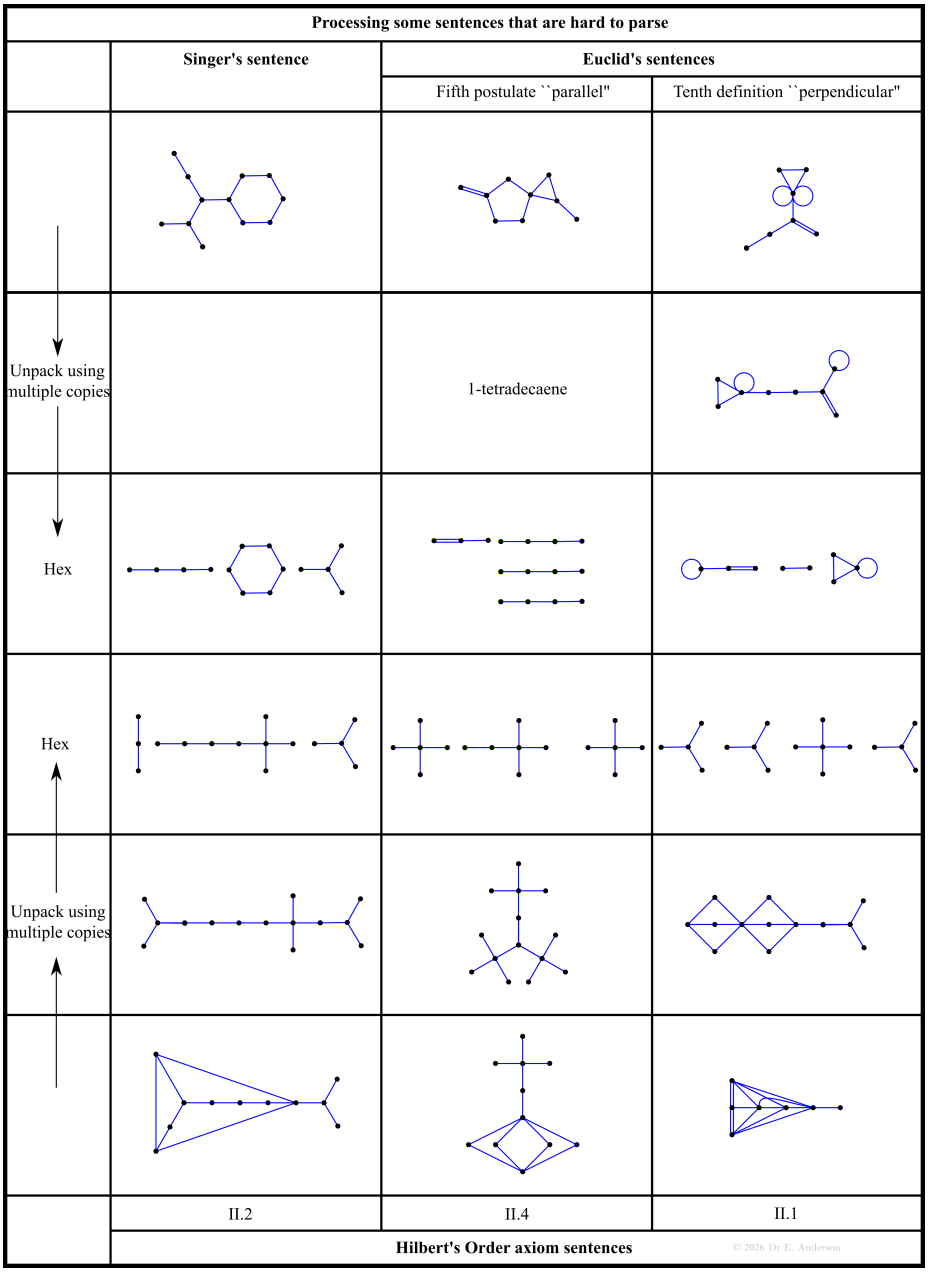

Figure 1:

[1]Institute for the Theory of STEM Date stamp: v1: 31-10-2026. Copyright of Dr E. Anderson. Cite as E. Anderson, “Syntax for clear Scientific exposition. 0. Preliminary Discussion", Online Encyclopaedia of Applied Grpah and Order Theory institute-theory-stem.org/oeagot-syntax-0 (2026).

1 Introduction: sentence word- versus concept-counts

Socratic Sage (SS): Today I shall start to teach you how to write clear Scientific exposition. Before to starting to wrestle with this, do you have any prior yoga to declare in this regard?

Pupil 1 (MP: Mathematics and Physics): George Ellis advised me to write 20 or less words per sentence.

SS: And you implement this how?

MP: I standardize my typing screen to 155-character gauge. If any sentence comes close to not fitting on one line, then I redraft it as two sentences...

♠ ◇ ♣

Pupil 2 (CM: Theoretical Computer Science and Mathematics): Your infrastructure is that LaTeX boxes are a coding environment, and so come with a monospaced font as default...

SS:

$$\frac{155 \ - \ 20}{20} \quad = \quad 6.75 \quad \text{is somewhat larger than} \quad (\,\text{average wordlength}\,)\,.$$

Which is around 4.65 in English, or 5.6 in Scientific text. Account for this!

MP: 155 characters jointly accommodates my formula environment. Also my effective average word-length is slightly longer in characters, due to containing tagged LaTeX macros. Finally, I did say “close to"...

SS: Fine. And your redrafting is certain to be hexed by thereby-whiches. But this is a good start... Today I shall teach you how to write using “Geometry: Plane and Fancy" [3]. For clear Scientific exposition, sentence word-count is majored by the following. Scientific concept-count in each sentence! Count these out for Singer’s opening portion...

♠ ◇ ♣

MP and CM, alternating: 7, 3, 2, 4, 3, 2, 3, 2, 2 .

SS: And now for the first Euclid quote.

MP and CM alternating again: 2, 3, 3, 4 .

SS: And now Singer again!

MP and CM alternando: 2, 5, 4, 4, 4, 6, 2, 4, 4, 12, 4, 3, 7, 3, 2, 3, 4, 4, 3, 1 .

SS: You know the drill...

MP and CM alternando: The translated Euclid quote has 5, 5, 5, 4, 2 .

Singer then wrote 2, 2, 2, 4, 6, 6, 4, 2 .

Back to Euclid, 2, 5, 5, 3, 5, 13 . ...

SS: This one is the famous fifth postulate on *parallelism*. Which is said by Singer to take around three times as many words as the others. This does not however necessitate cramming all of them into one sentence! Proceed!

MP and CM alternando: Singer wrote 3, 4, 7, 5, 4, 2, 2, 2, 3 .

Back to Euclid, 13 .

And then back to Singer once more: 4, 3, 3, 3, 4, 6, 2 .

♠ ◇ ♣

SS: This 13 is Euclid's tenth definition, of *perpendicularity*. Let us now skip ahead to Hilbert's quotations, and nothing but. So as to bring in a third author...

MP and CP alternando:

On *independence*, 5, 6, 3, 5 .

On *order*, 7, 7, 7, 12, 9, 4 .

[In the process, SS clarified that: person names are not Mathematical concepts. Attributed significance and Rhetorical elements are not either. The counts displayed are the corrected ones in these regards.]

♠ ◇ ♣

SS: Any observations so far?

MP: Singer would fit in my gauge, with one exception. Euclid and Hilbert however exhibit a significant proportion of sentences that I would flag to split up...

SS: Insert ". Which", ". By which". And then alternate ". By which" and ". Thereby"... How might we split Singer's large sentence?

MP: "The third definition can only be understood if we use the word "line" to refer to line segments. Which begin at a point A and end at a point B . And to complete lines, which have no ends or "extremities".

I.e. 4, 6, 4 . Hold on,

$$4 + 6 + 4 = 14 > 12 \ldots$$

So the splits do not conserve word count?

SS: Of course they do not. One part of the thereby-which's hex is that "X, which" counts as one concept. But "X. Which" counts as placing one copy of that concept in each of the two new sentences created...

[Some advice was also given at this point about treating such as "straight line" and "right angle" as priorly established composite single concepts.]

CM, split Euclid's large parallel and perpendicular sentences!

CM: Suppose that a straight line that falls on two straight lines has the following property. *It* makes the interior angle on the same side less than two right angles. *Then* the two straight lines, if produced indefinitely, meet in the following place. On that side on which the angles are less than the two right angles.

That is 3, 4, 4, 4 .

Suppose that a straight line set up on a straight line makes the adjacent angles equal to each other. Then each of the equal angles is right. And the straight line standing on the other is called a perpendicular to that on which it stands.

I.e. 3, 2, 3 .

SS: MP, take down Hilbert's between-point existence and lines triangle entry-exit symmetry axioms. I.e. axioms of order 2 and 4...

MP: *Let* A, C be points. Then there always exists at least one point B on the line AC with the following propertiy. That C lies between A and B .

I.e. 3, 7, 4 .

Let A, B, C be three points that do not lie on a line. Also let a be a line on the plane A B C . Which does not meet any of the points A, B, C .

That is 5, 6, 5 .

I can furthermore refine the first of these partly using non-sentences.

"Then there always exists at least one point B . On the line A C . With the following property".

Giving the split P(1) , P(2), P(3), S . See Fig 8.b).

SS: CM, wrap up the remaining long Hilbert sentence! That is the nature-of-betweenness axiom of order 1 ...

CM: Let A, B, C be points. Such that B lies between A and C . Then A, B, C are distinct points of a line. And B then also lies between C and A . "

That is, 4, 4, 5, 4 .

2 Approximate modelling by simple graphs

SS: I am satisfied that you two can now systematically Ellis your sentence lengths. I am however less satisfied with some of Hilbert's other sentences here than with the remainder of Euclid's. For they are denser in relators between concepts. And many Readers will not recollect seeing these shapes of sentence before. Or elsewise have trouble parsing them...

Let me put it to you that to first approximation, concept sentences are simple graphs [2, 4]. The concepts are the vertices and the relators are the edges.

One counting principle suitable for primary school-children is to never have more than 3 concepts in a sentence. And to also split up any occurrence of X implies Y implies Z implies X . So that the sentence simple graphs are always the 1- 2-path or 3-path! These are sentences that they will have heard many of before! And all of them can be parsed, in the manner of a very short chain...

For secondary-school children, the 3-cycle and the 4-graphs would be allowed. For university work, the 5-graphs. And in formulae rather than in the text environment, some 6-graphs.

But graphs dense in relators are to be excluded. Which 4-graphs might I not wish to see as sentence structures for secondary-school children, CM?

CM : Most of all, the complete graph K_4 . The next candidate to drop would be the diamond...

SS: Let us keep double-irreducible [6, 7] graph names as short and distinctive as possible. Tet and Di ! MP, which 5-graphs might be inadviseable in the text environment for university work?

MP: The complete pentagon. The same minus one edge. Wheel, Scotshouse, Scotsbanner, Spiderhead, Gem . With Pendant, Dart, Bowtie, House and $K_{3,2}$ as edge cases...

[These sentences neatly group the 10, 9, 8, 7 and 6-edge graphs on 5 vertices. See Fig 2.]

CM: Butterfly is a truer name than Bowtie ! For isomorphisms distort the Bowtie. But still recognizably flap the wings of a Cartoon Butterfly!

SS: Also use Tflap, Tleaf and Tarantulahead for the fourth to sixth please! For they are the flap, the leaf and the first homeomorph of Tet ... And not House and $K_{3,2}$ but Domicile and Didot for the first homeomorphs of Di ! Finally let Dant partner Dart as first leaves...

♠ ◇ ♣

SS: But why did you select these graphs?

MP: Maximize the edges, then strip edges away. All the while sticking to connected cases...

SS: How about a constructive principle?

MP (after a while): Two cycles is a stretch. Three or more is even worse...

SS: What else might be unacceptable, Dendrophila?

CM: The Claw-of-Claws would have too many branching points... With the Scorpion and bigger dipoles like the Dog as edge cases.

SS: I knew I could count on you to find the second Graph-Theoretic simplicity rule... Mycroft my rules, MP!

♠ ◇ ♣

MP (dutifully): Let cycles and branching points be faults. Four strikes and you are out! With branching strength [5] weighting rather than just by branching point count...

SS: Essentially. And as a rule of thumb. Why no weighting symmetry with cycle length?

CM: Because homeomorphs [4] can be used to thread in equation numbers and citations!

SS: Partly so. Concatenations of first uses of technical adjectives as well...

Also, because of how composing cycles itself carries excess branching point strength, we encounter a dilemma.

A) Either we insist on just unicycles and acycles = trees.

B) Or we select 5 strikes and you are out as the minimally-illicit case. Leaving 4 strikes as the edge case, rather...

Dendrophila, cover us with edge-case and minimally-illicit HI trees in the sense of horn B)...

[This produced Appendix A.2. SS then conditioned this quite large forest by the counting criteria of 9-vertex illicitness for HI trees...]

Now let us sing the graph names of the sentences that we have so far counted!

♠ ◇ ♣

MP and CM alternating: Lobster, P_3, P_2, S_3, P_3, P_2, P_3, P_2, P_2 .

SS: Hold your Dark Horses. Let us first split the sentances into paragraphs. Please also use the double-irreducible notation P in place of P_2 . And S in place of S_3 !

MP: Lobster, P_3, P, S

CM: And then P(1), P, P(1), P, P .

[CM was the first Pupil to adopt homeomorph-irreducible, and then double-irreducible notation for sentence graphs.]

SS: Now continue, one paragraph each...

MP (taking the hint): P, P(1), P(1), Paw .

CM: P, P(3), P(2), P(2) .

MP: S, P(4), P, P(2), P(2) .

CM: Pan-Q homeomorph irreducible. Third homeomorph on cycle, first on cis leaf. P(2), P(1), Paw(1). P(1), P .

♠ ◇ ♣

SS: Which grimace we split into?

CM: P(2), S $\left(1^3\right)$, S !

MP: P(1), P(2), P(2) . P(1), D .

CM: P(3), S(1), S(1) . C(1), P .

MP: P, P, P, P(2) . S($1^2\,0$) . S(2), P(2), P .

CM: P, S(1), Pan(0; 1) . P(1), P(3) . Spectacles , illegal, illegal: 8-graph double-irreducible. 2 cycles and 2 branching points! Second homeomorph on one cycle and double edge on one prong! Which we split into Propene and three P(1) !

SS: Just so. But use Pan , not Paw . For then Banner is a widely used phonetic first homeomorph...

MP: P(1), P(2) . S $\left(1^3\,0\right)$. S(1), S, P, . P, P, P(1) .

♠ ◇ ♣

CM: Why is it always me? Pan-Claw double irreducible. 1-homeomorph on one prong. Double-edge on the other. 2 loops on the ankle!

SS: Well, I did say simple graphs *to first approximation*...[3] Sliced up how?

CM: Propene with opposite end loop , P and C with loop.

SS: Finish off the paragraph, so it is *not* always you...

CM: P(1), P(2) . P(1), P(1), P(1) . 1-Butene, Pan(1; 1), P !

MP: Di(1), Di(2) . P(1), P(4) . A leaf on a 6-graph double-irreducible. Which in turn is the complement of the union of the 3-, 2- and 1-paths. Definitely illicit! S, S $\left(1\,0^3\right)$, S . 2-tailed Scorpion homeomorph-irreducible with excess arm-lengths 4 and 1 ...

SS: Claw-Plus-Claw , please! To CM's dipole = higher-order 2-path, this is a 3-path. With middle element carrying excess branch-strength 1 . And what graph do you get if you identify copies of the same objects?

MP (after a pause): Tet-Claw homeomorph irreducible. With (3, 1) homeomorph like so. (Draws bottom of first column of Fig 1.)

SS: And what is the graph version of our previous split of this?

MP: For now P(1) S $\left(3\,0^3\right)$ S .

♠ ◇ ♣

SS: Dendrophila!

[2] The middle one is Lobster from the starlike graph perspective. As parametrized by excess radiance length. Q is itself the quadruped tree; see Fig 4 for each.

[3] These forbid both loops and multi-edges [5].

CM: Next we have a second homeomorph of ... Range-Plus . Such that its Range part becomes $K_{4,2}$. The other extra vertex is on the shaft. (See bottom of column 2 of Fig 1).

Split into Plus , Cross , Plus !

SS: And MP to swing low and carry me home!

MP: Lobsterplus , I mean $S\left(1^4\right)$. Then P(2) thrice. Finally, Pan(1) and the smallest asymmetric tree S(2 1 0) !

3 Conclusion

SS (beginning to grow): And so, is such clear Scientific exposition to be based on Human-Language Syntax or Computer-Language Syntax?

CM: I object! Semantics is more significant than Syntax!

SS: Were we Coding, or Debating, I would say Yes. But in the business of clear Scientific exposition, a rotund No. Why this is so, and in which arenas Syntax majors Semantics, and vice versa, I leave to you as Riddles!

MP: Human-Language Syntax suffers from the tyranny of the sentence...

CM: Whereas Coding Syntax is scale-invariant...

SS: I like where this is going...

♠ ◇ ♣

CM: At most 5 concepts per sentence. At most 5 sentences per paragraph. At most 5 paragaphs per section...

MP: The Illuminati Trilogy implies that 5 is too many for the Common Man...

SS: Size-5 double irreducibles, with 3 strikes is out, will do for the Theory of STEM. The Everyday Person only went to school, however. For which I posited size-4 ...

CM: Scale-invariant Computer-Language Syntax for clear Scientific exposition!

SS: By the way, your tyrant's writ belongs on the bottom floor. An opposing denizen of the top floor is the essay-plan spider diagram... And yet this one readily hierachically extends all the way top-down. Indeed, what we have prescribed here is the triumph of parseable hierarchical 'spider diagram' Computer-Language Linguistic Syntax. Over the bottom-up expansion of Human-Language Linguistics from sentence analysis to texts of all lengths...

♠ ◇ ♣

SS: You will now find that 40-draft articles are a thing of the past, Chronologos. One draft for sentences will do. Then one for paragraphs... For another part of the thereby-which's hex is that "X, which" to "X. Which" both changes a paragraph's sentence count. And imparts the relator "share a concept with concatentation" between the two new sentences created. With iteration at each level of the hierarchy. Returning

$$O(\log N)$$

drafts for a work of length N .

Halmos'

$$O\left(N^2\right)$$

bound on an N-chapter book may moreover still apply.

CM: What is Halmos' bound?

SS: Upon writing chapter 2, rewrite chapter 1 to fit. Upon writing chapter 3, rewrite chapters 2 and 1 to fit [1] ... And so on, child-Gauss summing to

$$\binom{N}{2} = O\left(N^2\right) .$$

Halmos furthermore opines that this bound is usually far from tight!
These are conceptual-upgrade backwriting-redrafts. Partly Semantic: what to emphasize, consistentizing the motivation... Partly a matter of figuring out exactly which definitions, examples and lemmas subsequent text needs to sit upon. And partly a matter of evolution of how one layers the levels of Mathematical structure that the book turns out to need. As opposed to our parseable Syntax shell...

Acknowledgments Elsewise participants in this 2018 discussion expressed their right not be coauthors of this preamble. With thanks also to George Ellis for a previous discussion with one of the participants. And to Paul Halmos for being Paul Halmos.

A The smallest simple graphs

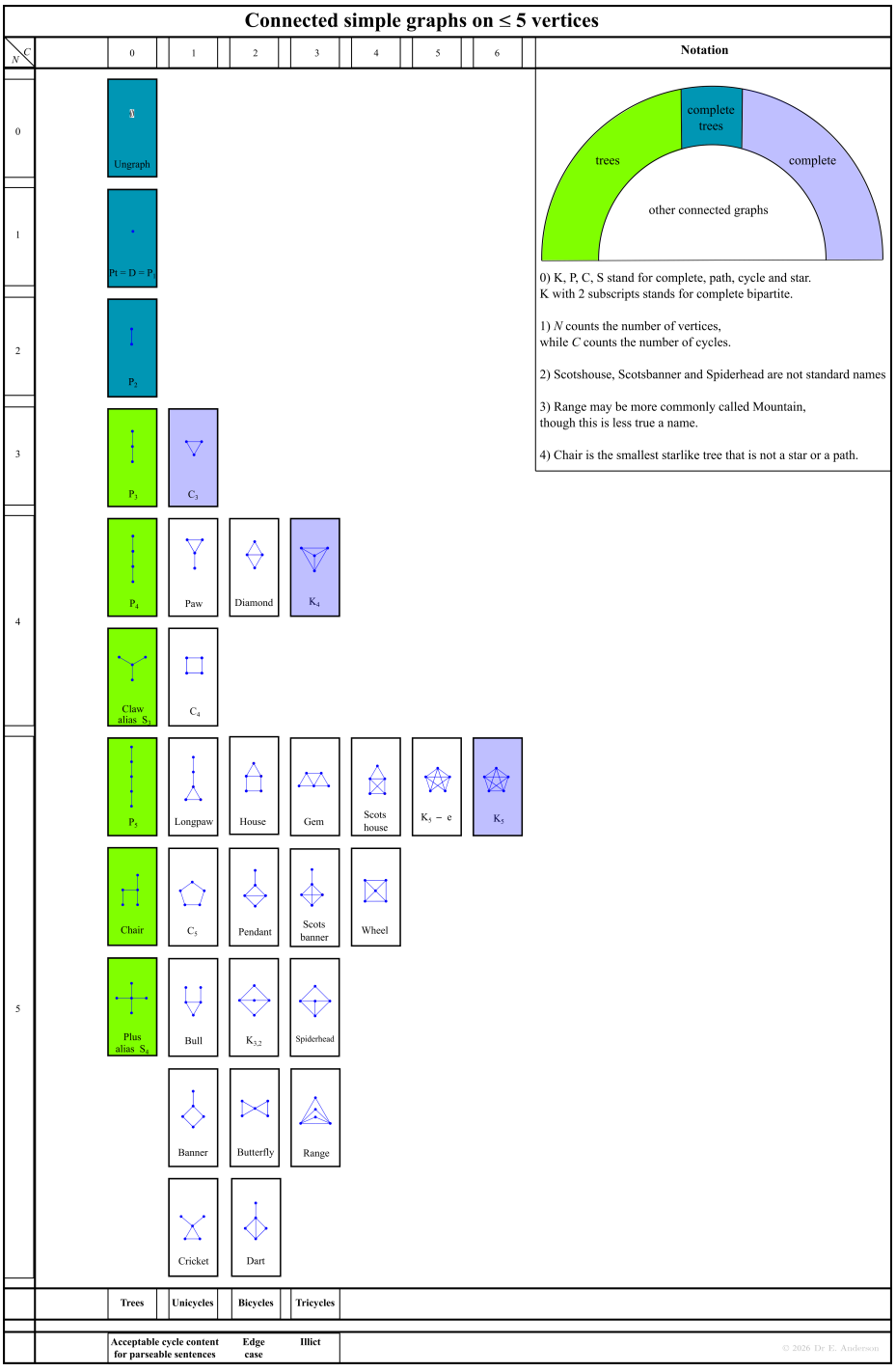

Figure 2:

Remark 1 Here $C = C(\mathrm{G})$ [5] is the *covering number by elementary cycles* of the graph G .

A.1 The smallest simple-graph homeomorph irreducibles

Remark 2 We next entertain a more Topological notion of simple graph...

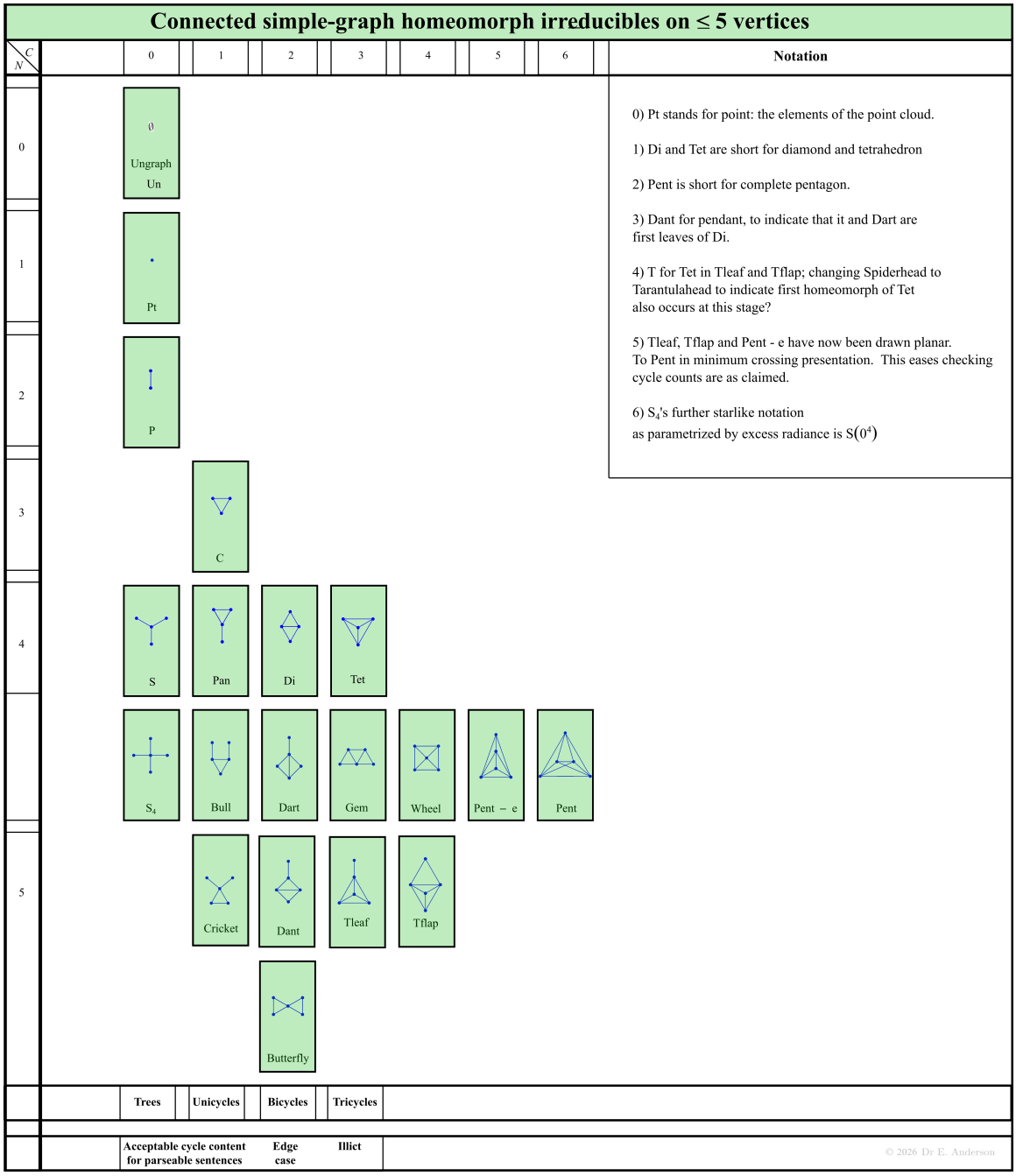

Figure 3:

A.2 Dendrophila's grove: trees mentioned in current text

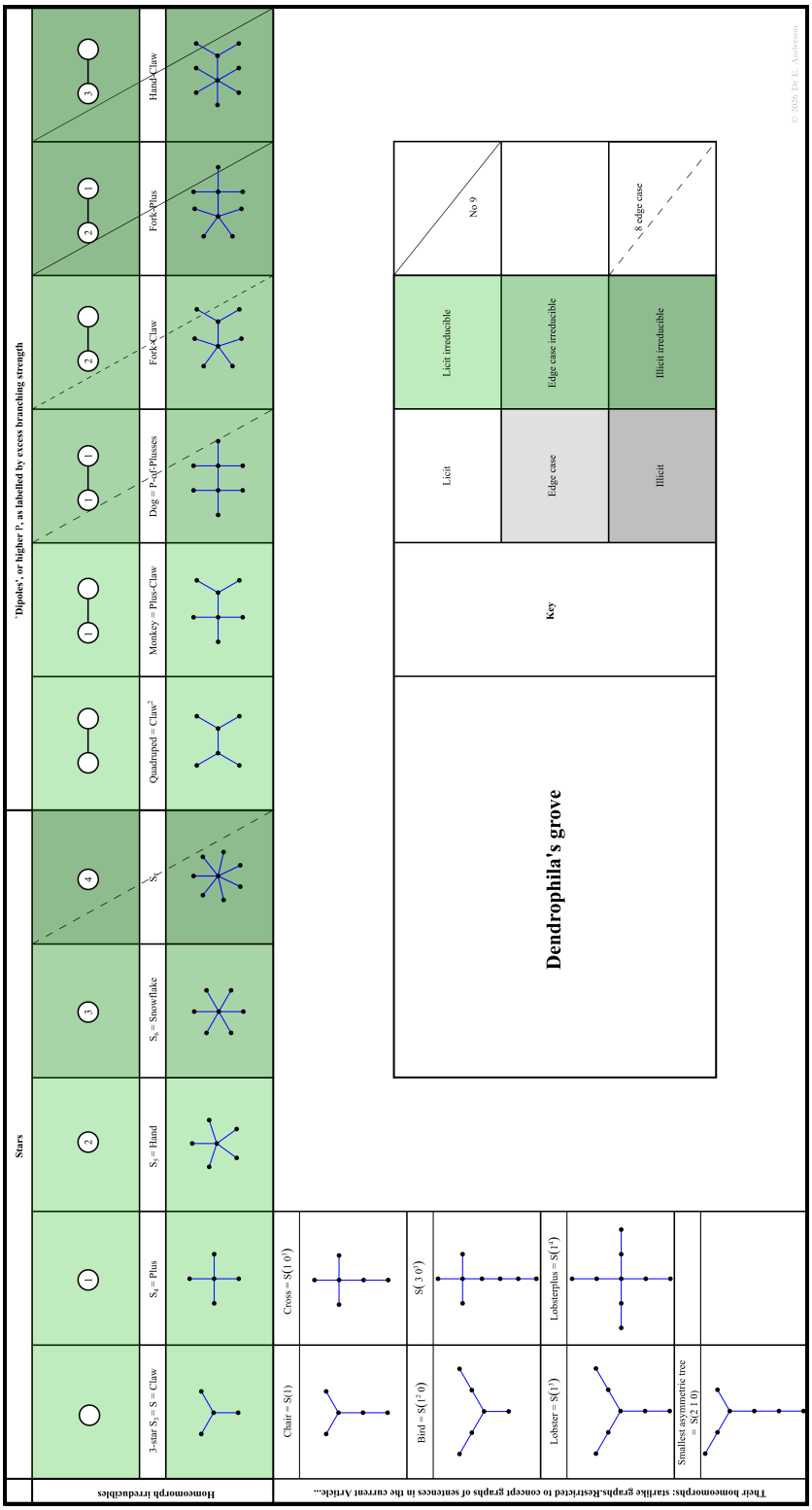

Figure 4:

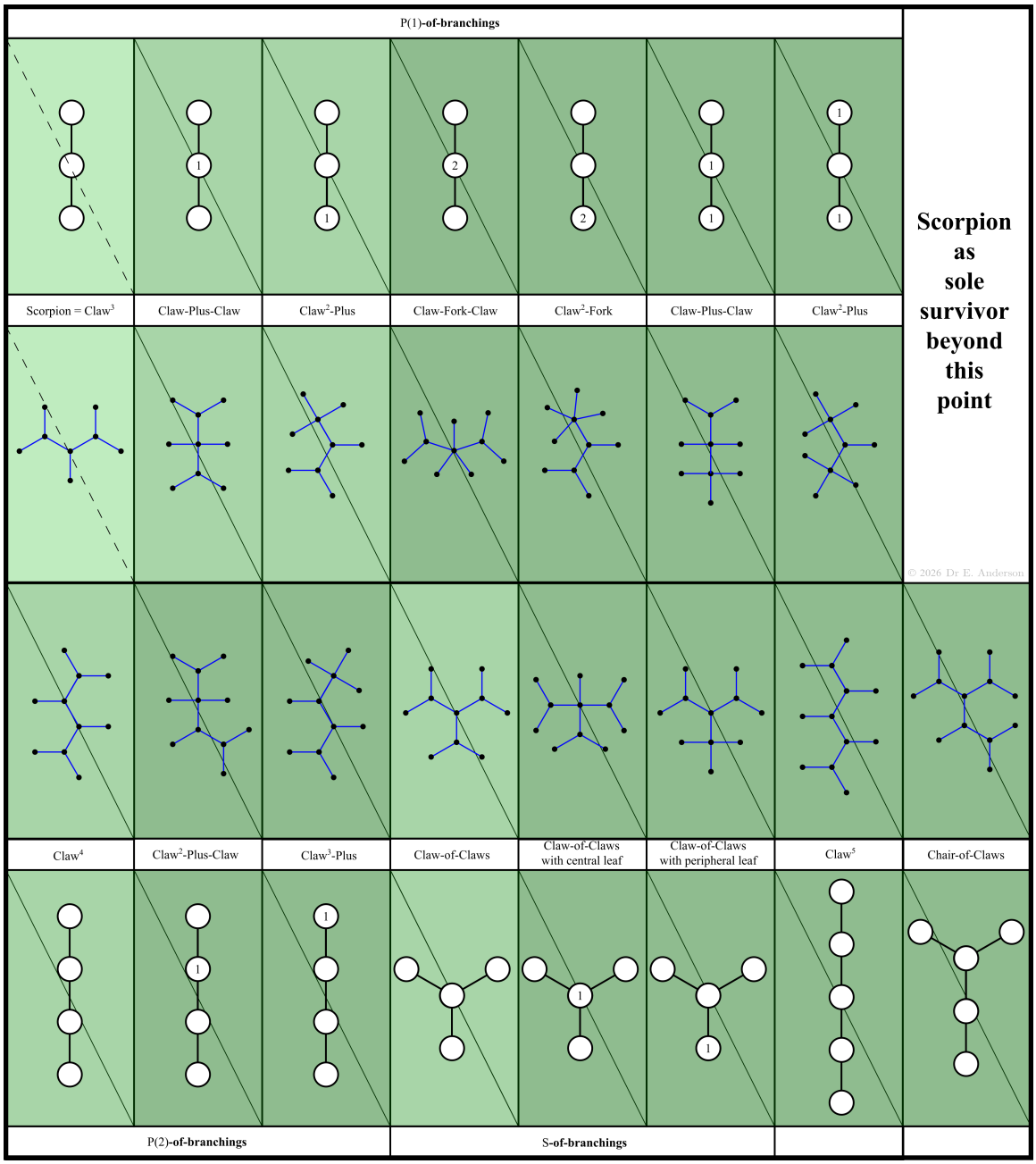
P(1)-of-branchings
Scorpion as sole survivor beyond this point
Scorpion = Claw3
Claw-Plus-Claw
Claw2-Plus
Claw-Fork-Claw
Claw2-Fork
Claw-Plus-Claw
Claw2-Plus
© 2026 Dr E. Anderson
Claw4
Claw2-Plus-Claw
Claw3-Plus
Claw-of-Claws
Claw-of-Claws with central leaf
Claw-of-Claws with peripheral leaf
Claw5
Chair-of-Claws
P(2)-of-branchings
S-of-branchings

Figure 5:

A.3 The smallest simple-graph double irreducibles

Remark 1 Now we defoliate [5, 6, 7] the graph as well...

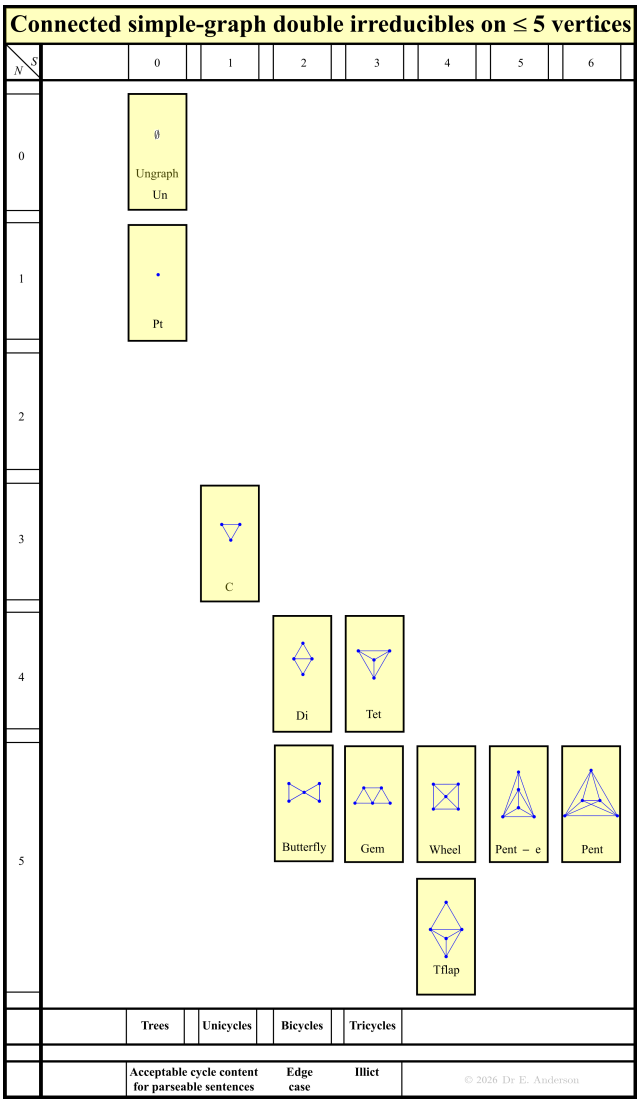

Figure 6:

Remark 2 Let us finally pass to parametrizing by SS's adapted variable

$$S \ := \ C \ + \ Br \ . \tag{1}$$

For *branching-strength of a graph*

$$Br(\mathrm{G}) \ := \sum_{v \ \in \ \mathfrak{V}(\mathrm{G})} br(v) \ . \tag{2}$$

Where in turn the *excess branching strength* of a vertex v is given by the following.

$$br(v) \ = \ \max(\, \mathrm{d}(v) \ - \ 2, \, 0 \,) \ . \tag{3}$$

For $\mathrm{d}(v)$ the degree [2, 4, 5], alias valency, of vertex v .

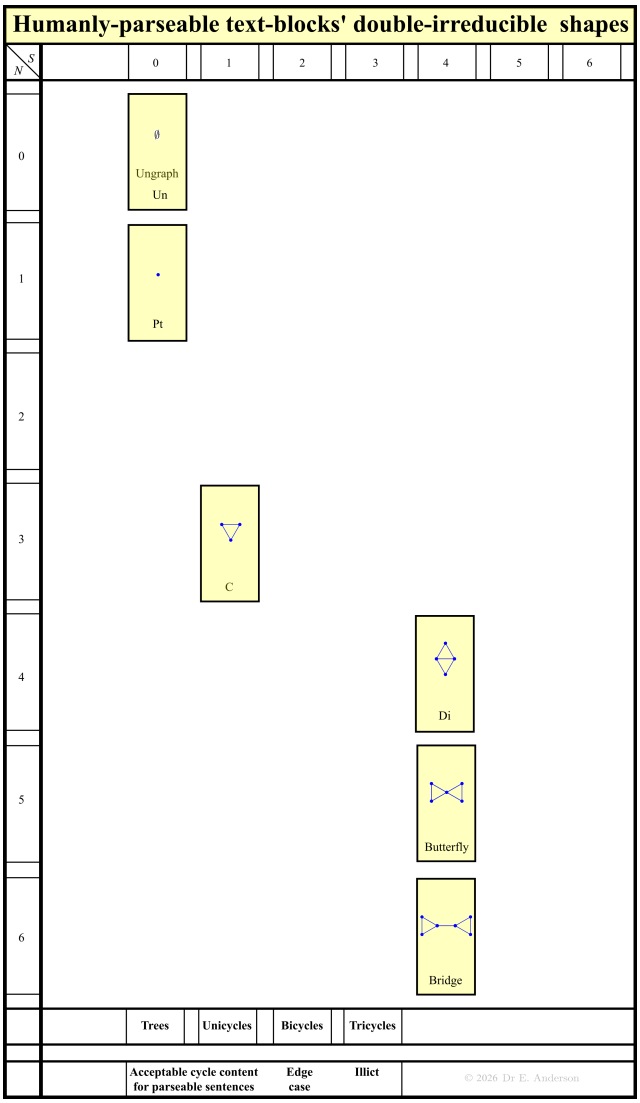

Figure 7:

Remark 3

$S = 4$ is the current Article's final conceptualization of 'edge case'.

$S = 5$ then returns the minimumly-illicit cases.

Dendrophila's grove's superposed cloud and grey backgrounds correspond to these cases. To extend this analysis from double-irreducibles and trees to homeomorph-irreducibles and all simple graphs, peruse Article 2.

Remark 4 Furthermore Butterfly and Bridge present an obvious choice of vertex to hex and edge to cut respectively. Sending these to pairs of smaller sentences whenever necessary. Splitting up Di is furtherly wasteful in duplication.

A.4 Other graphs mentioned in the current Article

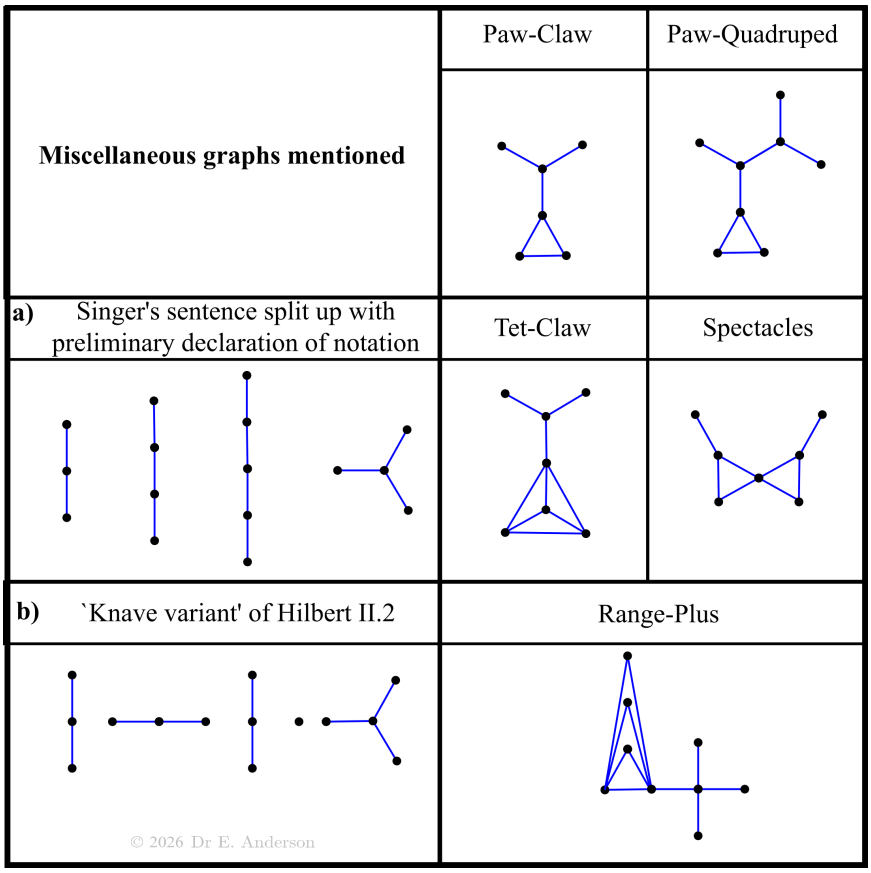

Figure 8:

Afterword SS: On the one hand, Singer, Euclid and Hilbert's interesting sentences feature as graphs on the cover figure of this account. On the other hand, the whole point of clear Scientific exposition is not to have any interesting sentences. Or paragraphs, or sections, and so on hierarchically.

The cover figure is there to illustrate a process. By which any interesting sentences you write or find can be processed into parseable sentence graphs. Thereby, the whole point of clear Scientific exposition is a close parallel to complex integration. Where one is to select a contour that splits into arcs, so that none of these require evaluating any hard integrals!

References

[1] P.R. Halmos, *How to Write Mathematics* L' Enseignement Mathématique **16** 9 (1970).

[2] R.J. Wilson, *Introduction to Graph Theory* (Longman, Edinburgh 2010). The first edition dates to 1972.

[3] D.A. Singer, *Geometry: Plane and Fancy* (Springer-Verlag New York 1998).

[4] D.B. West, *Introduction to Graph Theory* (Prentice–Hall, Upper Saddle River N.J. 2001).

[5] E. Anderson, *Applied Combinatorics*, Made freely available in response to the pandemic here: institute-theory-stem.org/combinatorics/ .

[6] E. Anderson and A. Ford, Online Encyclopaedia of Applied Graph and Order Theory, "Simple Graphs' 8 Double-Irreducibility Classes", institute-theory-stem.org/oeagot-graph-double-irreducibility-classes/ (2026).

[7] E. Anderson, Online Encyclopaedia of Applied Graph and Order Theory, "Smallest Nontrivial Arenas of Double-Irreducible Simple Graphs", institute-theory-stem.org/oeagot-graph-smallest-double-irreducible-arenas/ (2026).