



A Universal Grammar for Specifying Visualization Types

Yuri Engelhardt¹(✉)  and Clive Richards² 

¹ University of Twente, Enschede, The Netherlands
yuri.engelhardt@utwente.nl

² Birmingham City University, Birmingham, UK
clive.j.richards@me.com

Abstract. A ‘universal grammar’ for the full spectrum of visualization types is discussed. The grammar enables the analysis of any type of visualization regarding its syntactic constituents, such as the types of visual encodings and visual components that are used. Such an analysis of a type of visualization, describing its compositional syntax, can be represented as a specification tree. Colour coded tree branches between constituent types enforce the combination rules visually. We discuss how these specification trees differ from linguistic parse trees, and how visual statements differ from verbal statements. The grammar offers a basis for generating visualization options, and the potential for formalization and for machine-readable specifications. This may serve as a basis for a system providing computer-generated visualization advice.

Keywords: Visualization types · Visual encodings · Syntax of diagrams · Specification trees of visualization types · Visually enforced syntax rules · Comparing visualization types · Comparing verbal and visual statements

1 Purpose of This Work

The grammar presented here is the most recent addition to the ‘DNA of visualization’, a framework that may help designers to generate visualization options. Descriptions of this framework can be found in Engelhardt and Richards 2018, Richards and Engelhardt 2020, and at [VisDNA.com](https://visdna.com). The framework also offers a tool for research, a basis for formalization, and the potential for computer-based visualization advice. Put another way, by defining the fundamental building blocks of visualization, their interrelationships and the grammar for their combination, as discussed in this paper, the framework provides a method for deconstructing visualizations – which in turn provides a toolkit for exploring design choices. The system may even support combinations of visual encoding possibilities that result in entirely novel visualization types. The ‘DNA of visualization’ can thus be thought of as a *compositional* taxonomy of visualization, which goes beyond what other taxonomies offer. This work can also be framed as a *pattern language* for

Y. Engelhardt and C. Richards—Both authors contributed equally to the work.

© The Author(s) 2021

A. Basu et al. (Eds.): Diagrams 2021, LNAI 12909, pp. 395–403, 2021.

https://doi.org/10.1007/978-3-030-86062-2_40

visualization, or as a system to describe ‘the structure of the information visualization design space’ – in the sense of the paper with that title by Card and Mackinlay (1997, IEEE VIS ‘Test of Time Award’ in 2017).

We use both a linguistic analogy of ‘parts of speech’ and ‘grammar’ (see Sects. 2 and 5), and the biological metaphor of ‘DNA’ and ‘species’ (see Sect. 3) when discussing types of visualizations, and commonalities and differences between them.

Like academic work in linguistics, the work presented here is primarily not prescriptive but descriptive, in the sense that it facilitates the understanding, modelling and creation of (visual) language.

2 A Grammar for ‘Parts of Graphical Speech’

We share with Fred Lakin the view that, “When a person employs text and graphic objects in communication, those objects have meaning under a system of interpretation, or ‘visual language’.” (Lakin 1987, p. 683). The constituents of visualizations – which include *visual components* as well as *visual encodings* – can be conceived of as “parts of graphical speech”, in the sense that Graham Wills is suggesting: “a visualization can be defined by a collection of ‘parts of graphical speech’, so a well-formed visualization will have a structure, but within that structure you are free to substitute a variety of different items for each part of speech” (Wills 2012, p. 22).

We offer a ‘universal grammar’ that describes how the constituents of a visualization – Graham Wills’ ‘parts of graphical speech’ – can be combined into a specification tree¹ describing that type of visualization. This grammar is expressed visually, through colour-coded couplings between constituents, enforcing the combination rules in a visual way (see Sect. 6 and Fig. 1). This facilitates the systematic and detailed analysis of commonalities and differences between one type of visualization and another, as well as the exploration of visualization options.

This system can be applied to the full diverse spectrum of different types of visualization. While Wilkinson’s ‘grammar of graphics’ (2005) or the *Vega-Lite* visualization grammar (Satyanarayan et al. 2016) can be used to describe many statistical visualizations, these frameworks are unable to deal with, for example, most non-statistical visualizations. The framework presented here covers statistical visualizations as well as non-statistical visualizations, such as family trees, Venn diagrams, flow charts, texts using indenting, technical drawings and scientific illustrations.

3 ‘DNA’ and ‘Species’ – A Metaphor for Visualizations

In their ‘Tour through the Visualization Zoo’, Jeffrey Heer et al. (2010) say that “all visualizations share a common ‘DNA’ – a set of mappings between data properties and visual attributes such as position, size, shape, and color – and that customized species of visualization might always be constructed by varying these encodings.” (ibid. p. 60). We use this metaphorical idea of the ‘DNA of visualization’ in a similar vein, taking

¹ Engelhardt (2006) proposed a grammar-driven analysis of graphics with tree diagrams. The paper included an earlier version of the specification tree shown here in Fig. 1.

it to the extent of identifying a comprehensive set of individual DNA building blocks of visualization, and the rules for combining them. This allows for the construction of a broad range of different types of visualizations – Heer’s “customized species of visualization”. We will refer to the DNA of visualization as ‘VisDNA’. VisDNA building blocks are shown as colour coded three-letter abbreviations, see Fig. 1.

We offer a grammar for combining VisDNA building blocks. The grammar rules are presented in Sect. 5 and at VisDNA.com. We refer to a ‘well-formed’ combination of VisDNA building blocks, i.e. one that follows the rules, as a **visualization species**. Many common *visualization species* have been given a name (e.g. ‘pie chart’) and are generally referred to as ‘chart types’, while novel or rare visualization species often do not have a name (yet). As Heer et al. (2010, p. 67) write, “many more species of visualization exist in the wild, and others await discovery.” We have analyzed a large number of *visualization species* using our system, including most of the corpus at datavizproject.com plus many other examples. Examples are shown in Fig. 1 and Fig. 2. Many more examples can be found at VisDNA.com.

4 Visual Encodings: Arranging, Varying, Linking

Visual encodings are at the centre of the VisDNA system (other building blocks include *types of information*, *visual components*, *layout principles* and *directions* – see Richards and Engelhardt 2020, and our accompanying website VisDNA.com). *Visual encodings* can be divided into three subgroups. By applying *visual encodings*, visual components can be spatially *arranged* in order to construct visualizations, *var-ied* regarding their visual properties, or *lin-king* by adding *configurator components*.

- **Arranging:** All the *types of information* (grey DNA) in the VisDNA system can be represented by how visual components are spatially *arranged* into a meaningful configuration. Examples of *arranging* are positional encodings such as *grouping by position*, *positioning on an axis*, *nesting*, or *coupling by adjacency* (red DNA).
- **Varying:** Quantity, order and category membership can be represented by how visual components are visually *var-ied*. Examples of *var-ying* are visual encodings such as *colour coding* or *sizing* (blue DNA).
- **Linking:** Relationships between entities, and in some cases category membership, can be represented by *lin-king* visual components using *configurator components* (such as connector lines or boundaries). These visual encodings are *connecting* and *grouping by boundary* (pink DNA).

A visual component can be involved in *several different* visual encodings, simultaneously representing different types of information.

5 A Visual Grammar for Combining VisDNA Building Blocks

We are proposing a visual grammar for combining VisDNA building blocks, the key rules of which are given in Fig. 1. This grammar includes colour-coded representations of VisDNA building blocks and their couplings. The colour codings enforce the key

The visual grammar of VisDNA: colour-coded couplings enforce the combination rules – only matching colours can connect

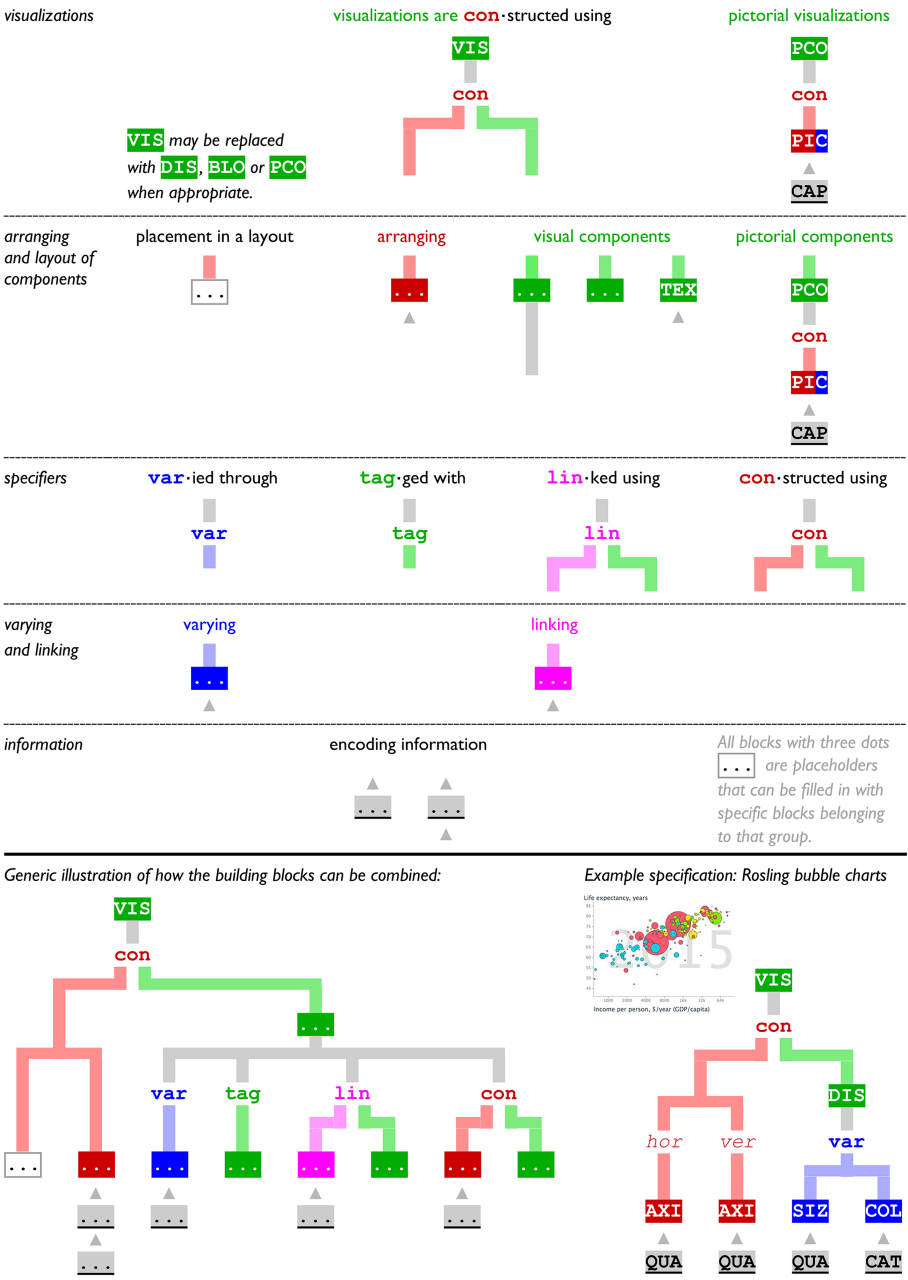


Fig. 1. The visually enforced key grammar rules for combining VisDNA building blocks, and an example visualization with the VisDNA specification tree that defines this visualization species. (Color figure online)

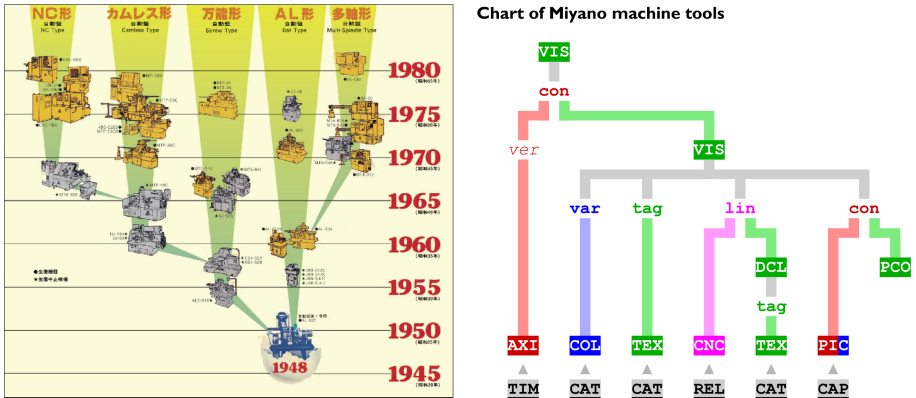


Fig. 2. Example visualization (courtesy of Citizen Machinery Miyano Co. Ltd.) with the VisDNA specification tree that defines this visualization species. (Color figure online)

combination rules in a *visual* way – only couplings of matching colours can connect, forming the branches of a VisDNA specification tree.

Various further rules for combining VisDNA, for example regarding the constraints for connecting visual encodings to types of information, to directions, and to visual components, are given at VisDNA.com.

A VisDNA specification tree specifies a *visualization species*. All individual specimens belonging to the same visualization species can thus be described by the same VisDNA specification tree. A visualization species can be transformed into another visualization species by adding, replacing or removing one or more VisDNA building blocks.

6 VisDNA Specification Trees

VisDNA specification trees are drawn so that they have a layer at the bottom showing the *types of information* (grey DNA) that are represented. The layer directly above that includes the *visual encodings* (red/blue/pink DNA) that are used to represent those types of information, plus any *layout principles* (black-on-white DNA) that may be involved. The remaining layers above the visual encodings show the **specifiers** ‘*var-ied* through’, *tag-ged*² with, ‘*lin-ke*d using’ and/or ‘*con-struct*ed using’ (in that left-to-right order). These *specifiers* characterize *visual components* (green DNA) by connecting them to the *visual encodings* in which they are involved, and when needed, to other visual components (either subcomponents, configurator components or tags), integrating any *directions* that may apply. The branches come together in a single node at the top.

² *Tagging* refers to the identification or annotation of visual components with either text, symbols or embedded visualizations. Tagging is a feature of most visual representations of information.

7 Verbal Statements and Visual Statements

We have found it useful to borrow from linguistics the idea of a ‘grammar’ and of analyzing representations using tree diagrams, applying these to visualizations. It is clear though that visualizations differ from expressions in verbal languages in a number of ways.

In verbal languages, a statement is constructed from a number of words and their *sequential order* within a sentence. A single word, on its own, usually does not represent a statement. Making an analogy with visualization, one may think that *words* could correspond to *visual components* within a visual representation.

However, a visual component can operate at a higher level than a single word. In a visual representation, a single visual component (green DNA in our framework – including, if present, its label) can represent a statement through its involvement in a *visual encoding*. This may include its position in a meaningful *spatial arrangement* (usually non-sequential – red DNA), its *visual properties* (blue DNA), and any relationship to a *configurator component* (e.g. to a *boundary* or to a *connector line* – pink DNA). For example, a symbol together with its colour gradient may represent the statement ‘this measuring station records a *medium* level of pollution’.

A single visual component can even be involved in *several* visual encodings simultaneously (regarding the component’s position, its visual properties, and its relationships to configurator components), thus representing *several* statements through a *single* visual component.

Let us consider, for example, a dot in a scatter plot. Such a single dot may represent that ‘the UK has an average life expectancy of 81 years and a GDP per capita of 40,000 dollars’. This single dot thus makes one or more statements. To take another example, a single symbol within the intersecting circles of a Venn diagram may represent that ‘dolphins are mammals, and not fish’. Every single symbol added within the circles of a Venn diagram thus makes additional separate statements. As a final example, a bar in a bar chart may represent that ‘2 cm of rain fell on April 1st’. Every single bar added to the bar chart likewise makes an additional separate statement.

In summary, in verbal languages, in order to represent a statement, one can combine words into a sentence. In visualizations, in order to represent one or more statements, one can apply one or more *visual encodings* to a *single* visual component. A visualization usually contains a number of visual components that are following the same visual encoding rules, and thus visualizations represent *sets* of statements – statements that are characterized by the same syntactic structure.

8 VisDNA Specification Trees and Linguistic Parse Trees

For spoken and written languages, linguistic frameworks have been developed that enable the parsing of expressions. Parsing a linguistic expression is based on a grammar of the concerned language, and involves dividing an expression into its (syntactic) constituents and the identification of the (syntactic) relationships between these constituents. The result of parsing an expression can be represented in a parse tree. The development of grammars and parsing are regarded as key accomplishments of the field of linguistics.

Syntactic constituents of a sentence are also referred to as ‘parts of speech’. At the beginning of this paper we quoted Graham Wills: “a visualization can be defined by a collection of ‘parts of graphical speech’, so a well-formed visualization will have a structure, but within that structure you are free to substitute a variety of different items for each part of speech” (Wills 2012, p. 22).

Linguistic parse trees and VisDNA specification trees both identify syntactic constituents of a representation – ‘parts of speech’ or ‘parts of graphical speech’ – and the syntactic relationships between those constituents, thus describing the representation’s compositional syntax. This is what linguistic parse trees and VisDNA specification trees have in common – in other regards they are quite different.

In linguistics, a parse tree describes the syntactic categories of words and their combination into a *sequential* order in a specific sentence.

In visualization, due to its very nature, syntactic constituents include not only different types of *visual components*, but also *visual encodings* – different types of spatial *arranging*, *visual varying*, and *linking* with configurator components. A VisDNA specification tree describes the syntactic categories of such constituents of a visualization species and their *simultaneous* combination. For example, a symbol’s position in a visualization, its size and its colour, all exist simultaneously, rather than in a sequential order. If applicable, a VisDNA specification tree also indicates the nesting of smaller visual components within larger ones – which is also non-sequential.

Another difference is that, in linguistic analysis, usually every individual word is featured in a parse tree, while in VisDNA specification trees only sets of constituents are specified. The number of instances in each set (e.g. the number of *visual components*, or the number of colours used in a *colour coding*) is not part of the specification of a *visualization species*.

9 Future Work

This framework is an evolving programme. Because of the flexible structure of the framework, further types of VisDNA building blocks may be added, in order to accommodate any additional visualization species that one may want to describe and that cannot be fully specified using the current scheme. Examples may be the addition of VisDNA building blocks for animation or interactivity in visualizations.

VisDNA specification trees offer a potential research tool for exploring various kinds of commonalities, family resemblances and differences between visualization species within collections of a wide range of visual representations. For example, an application of the VisDNA framework to neural network system diagrams has been described in Marshall et al. (2021, in this volume).

Future work in using this framework to compare visualization species may lead to a better understanding of the structure of the visualization design space. The VisDNA building blocks and the grammar proposed here, also offer the potential for formalization and for machine-readable specifications. This may serve as a basis for a system providing computer-generated visualization advice, which could be linked to an application, such as, for example, a future version of the grammar-based tool *Vega-Lite*, in order to produce actual visualizations and variants of them.

Acknowledgements. The authors would like to thank Guy Clarke Marshall for a number of valuable discussions about the VisDNA framework, and three anonymous reviewers for their thorough and useful comments.

References

- Card, S.K., Mackinlay, J.: The structure of the information visualization design space. In: Proceedings of the 1997 IEEE Symposium on Information Visualization (InfoVis 1997), pp. 92–99 (1997). <https://doi.org/10.1109/INFVIS.1997.636792>
- Engelhardt, Y., Richards, C.: A framework for analyzing and designing diagrams and graphics. In: Chapman, P., Stapleton, G., Moktefi, A., Perez-Kriz, S., Bellucci, F. (eds.) Diagrams 2018. LNCS (LNAI), vol. 10871, pp. 201–209. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-91376-6_20
- Engelhardt, Y.: Objects and spaces: the visual language of graphics. In: Barker-Plummer, D., Cox, R., Swoboda, N. (eds.) Diagrams 2006. LNCS (LNAI), vol. 4045, pp. 104–108. Springer, Heidelberg (2006). https://doi.org/10.1007/11783183_13
- Heer, J., Bostock, M., Ogievetsky, V.: A tour through the visualization zoo. *Commun. ACM* **53**(6), 59–67 (2010). <https://doi.org/10.1145/1743546.1743567>
- Lakin, F.: Visual grammars for visual languages. In: Proceedings of the sixth National Conference on Artificial Intelligence (AAAI 1987), vol. 2, pp. 683–688 (1987). <https://dl.acm.org/doi/10.5555/1856740.1856793>
- Marshall, G.C., Jay, C., Freitas, A.: Understanding scholarly neural network system diagrams through application of VisDNA. In: Basu, A., et al. (Eds.): Diagrams 2021, LNAI, vol. 12909, pp. 1–17 (2021)
- Richards, C., Engelhardt, Y.: The DNA of information design for charts and diagrams. *Inf. Des. J.* **25**(3), 277–292 (2020). <https://doi.org/10.1075/idj.25.3.05ric>
- Satyanarayan, A., Moritz, D., Wongsuphasawat, K., Heer, J. Vega-Lite: A grammar of interactive graphics. *IEEE Trans. Vis. Comput. Graph.* **23**(1), 341–350 (2016). <https://doi.org/10.1109/TVCG.2016.2599030>, <https://vega.github.io>
- Wilkinson, L.: *The Grammar of Graphics*, 2nd edn. Springer, New York (2005). <https://doi.org/10.1007/0-387-28695-0>
- Wills, G.: *Visualizing Time*. Springer, New York (2012). <https://doi.org/10.1007/978-0-387-77907-2>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

