South American Journal of Logic
Vol. X, n. X, pp. 1–38, 20XX
ISSN: 2446-6719

Planar Heyting Algebras for Children

Eduardo Ochs

Abstract

This paper shows a way to interpret (propositional) intuitionistic logic *visually* using finite Planar Heyting Algebras ("ZHAs"), that are certain subsets of $\mathbb{Z}^2$. We also show the connection between ZHAs and the familiar semantics for IPL where the truth-values are open sets: the points of a ZHA H correspond to the open sets of a finite topological space $(P, \mathcal{O}_A(P))$, where the topology $\mathcal{O}_A(P)$ is the order topology on a 2-column graph (P, A). The logic of ZHAs is between classical and intuitionistic but different from both; there are some sentences that are intuitionistically false but that can't have countermodels in ZHAs — their countermodels would need three "columns" or more.

In a wider context these ZHAs are interesting because toposes of the form $\mathbf{Set}^{(P,A)}$ are one of the basic tools for doing "Topos Theory for Children", in the following sense. We can *define* "children" as people who think mathematically in a certain way — as *people who prefer to start from particular cases and finite examples that can be drawn explicitly, and only then generalize* — and we can define a method for working on a particular case (less abstract, "for children") and on a general case ("for adults") in parallel, using parallel diagrams with similar shapes; we have some ways of transfering knowledge from the general case to the particular case, *and back*. This method is sketched in the introduction.

Except for the introduction this paper is self-contained, and its title "Planar Heyting Algebras for Children" also has a second sense, different from the above: it can be read by students who have just taken a basic course on Discrete Mathematics — who are "children" in the sense that they don't have much mathematical maturity — and it prepares these students to read standard books on Logic that they would otherwise find a bit too abstract.

This paper is the first in a series of three. Categories and toposes only appear explicitly in the third one, that is about visualizing geometric morphisms, and at this moment the method of parallel diagrams has only been fully formalized for categorical diagrams. Behind the choices of finite examples and particular cases in this paper there is an *attempt* to adapt that method to areas outside Category Theory, but the precise details of how this is done are left for a future work.

Keywords: Heyting Algebras, Intuitionistic Logic, diagrammatic reasoning.

This paper is the first in a series of three. Let's refer to them as PH1 (this one), PH2 and PH3, and to the whole series as PH123. A nearly complete working draft of PH2 is available at [Och18], and the extended abstract [Och19b] shows the core results that will be in PH3.

The objective of the series can be explained in two ways. In the first one — shallow, and purely mathematical,

- PH1 shows how to interpret IPL in Planar Heyting Algebras ("ZHAs", sec.4) and shows that ZHAs are order topologies on two-column graphs ("2CGs", sec.14); this is used to show how one can develop visual intuition about IPL. The trickiest part is the implication; the method that allows one to calculate $P \to Q$ by sight in ZHAs has four subcases, and is discussed in sections 7, 8, and 9. It would probably be obvious to anyone who has worked enough with lattices, but I believe that it deserves to be more widely known.

- The paper PH2 extends the correspondence $(P, A) \leftrightsquigarrow H$ between 2CGs and ZHAs of PH1 to a correspondence $((P, A), Q) \leftrightsquigarrow (H, J)$ between 2CGs "with question marks" and ZHAs with a J-operator — that, more visually, are ZHAs "with slashings".

- PH3 transports this to Topos Theory: if we regard a 2CG (P, A) as a category, then $\mathbf{Set}^{(P,A)}$ is a topos whose objects are easy to draw, and the logic of $\mathbf{Set}^{(P,A)}$ is exactly the ZHA associated to (P, A); also, a set of question marks $Q \subseteq P$ induces an operation on $\mathbf{Set}^{(P,A)}$ that erases the information on Q and reconstructs it in a natural way, and this erasing-plus-reconstruction yields a sheafification functor — that is exactly the one associated to the local operator j associated to the J-operator J. This gives us a way to visualize (certain) toposes, sheaves, geometric morphisms, and two factorizations of geometric morphisms.

The second way to explain the goals of PH123 is by taking *Diagrammatic Reasoning* as the main theme. Let me start with an anecdote (90% true). Many, many years ago, when I tried to learn Topos Theory for the first time, mainly from [Joh77] and [Gol84], everything felt far too abstract: most of the diagrams were omitted, and the motivating examples were mentioned very briefly, if at all. The intended audience for those books surely knew how to supply by themselves the missing diagrams, examples, calculations, and details — but I didn't. My slogan became: "I need a version for children of this!".

At first this expression, "for children", was informal, and I used it as a half-joke. Very gradually it started to acquire a precise sense: clearly, CT done in a purely algebraic way is "for adults", and diagrams, particular cases, and finite examples are "for children". Writing "for adults" only and keeping the mentions to the "for children" part very brief is considered good style because "adults" have the technical machinery for producing more or less automatically the "for children" part when they need it, and people who are not yet "adults" can become "adults" by struggling with the texts "for adults" long enough and learning by themselves how to handle the new level of abstraction.

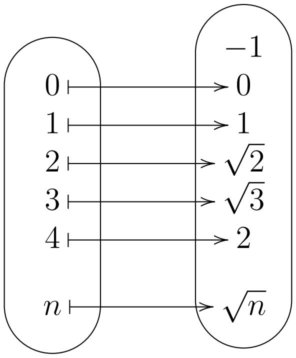
A clear frontier between "for adults" and "for children" appears when we realize that we can draw a diagram for the general case ("for adults") of a categorical concept and the diagram for a particular case of it ("for children") side by side. The two diagrams will have roughly the same shape, and we can transport knowledge between them in both ways: from the general to the particular, *and back*. Look at Figure 1; let's name its subdiagrams as A, B, and C, like this: $\substack{A\,B\\C}$. Each one of A, B, C has an *internal view* above and an *external view* below.

Diagram A shows, below, the external view of the function $\mathbb{N} \xrightarrow{\sqrt{\ }} \mathbb{R}$, and above that its internal view — in which one of the arrows, $n \mapsto \sqrt{n}$, shows the action of $\sqrt{\ }$ on a generic element, and the other '$\mapsto$' arrows, like $3 \mapsto \sqrt{3}$ and $4 \mapsto 2$, show substitution instances of $n \mapsto \sqrt{n}$, maybe after some term reductions.

Diagram B shows the external view of a (generic) adjunction $L \dashv R$, and above it its internal view. The nodes and arrows above $\mathbf{B}$ are objects and morphisms in $\mathbf{B}$, and similarly for the nodes and arrows above $\mathbf{A}$. The '$\mapsto$' arrows of the internal view are now of three kinds: actions of functors on objects, actions of functors on morphisms, and "transpositions" coming from the natural isomorphism $\mathrm{Hom}(L-,-) \leftrightarrow \mathrm{Hom}(-,R-)$. Diagram C is essentially the same as B, but for a particular adjunction: $(\times B) \dashv (B\to)$. Note how the diagrams B and C have exactly the same shape — but our diagrams for internal views are much bigger than the corresponding external views.

For a case in which the interplay between external and internal views is examined in full detail, see [Och19a]; it shows how each node and arrow in the diagrams can be can interpreted as a term in a type system, and this *may* be a basis for analyzing precisely which kinds of knowledge, and which kinds of intuitions — as in [Krö07], especially sec.1.3.2, and in [Cor04] — we are transporting from the less abstract diagrams to the more abstract ones, and vice-versa. Note that having a clearly-defined method for lifting information — in the sense of [Och13] — from a case "for children" to a case "for adults" would allow people to publish much more material "for children" than they do now, without this being regarded as bad style. For a non-trivial example of lifting information from a particular case to a general case, see [Och19b].

This paper can be seen as part of bigger projects in at least the two ways described above, but it was also written to be as readable and as self-contained as possible. In 2016 and 2017 I had the opportunity to test some of the ideas here on "real children", in the sense of "people with little mathematical *knowledge* and little mathetical *maturity*". I gave a seminar course about Logic and λ-calculus that had no prerequisites, and that



$$\begin{tikzcd}
\mathbb{N} \arrow[r, "\sqrt{\ }"] & \mathbb{R}
\end{tikzcd}$$

$$\begin{tikzcd}
& LA' \arrow[dd, "L\alpha"'] & & A' \arrow[ll, mapsto, "L_0"'] \arrow[dd, "\alpha"] & \\
& & \overset{L_1}{\longmapsfrom} & & \\
LRB \arrow[dd, "\epsilon B"'] & LA \arrow[dd, "\substack{g^\flat\\ f}"'] & & A \arrow[ll, mapsto, "L_0"'] \arrow[dd, "\substack{g\\ f^\sharp}"] & A \arrow[dd, "\eta A"] \\
& & \underset{\sharp}{\overset{\flat}{\substack{\longmapsfrom\\ \longmapsto}}} & & \\
B & B \arrow[rr, mapsto, "R_0"'] \arrow[dd, "\beta"'] & & RB \arrow[dd, "R\beta"] & RLA \\
& & \underset{R_1}{\longmapsto} & & \\
& B' \arrow[rr, mapsto, "R_0"'] & & RB' &
\end{tikzcd}$$

$$\begin{tikzcd}
\mathbf{B} \arrow[r, shift right, "R"'] & \mathbf{A} \arrow[l, shift right, "L"']
\end{tikzcd}$$

$$\begin{tikzcd}
& A{\times}C \arrow[dd, "\lambda p.(f(\pi p){,}\pi' p)"'] & & A \arrow[ll, mapsto] \arrow[dd, "f"] & \\
& & \longmapsfrom & & \\
(C{\to}D){\times}C \arrow[dd, "\lambda p.(\pi p)(\pi' p)"'] & B{\times}C \arrow[dd, "\substack{\lambda p.g(\pi p)(\pi' p)\\ h}"'] & & B \arrow[ll, mapsto] \arrow[dd, "\substack{g\\ \lambda b.\lambda c.h(b{,}c)}"] & B \arrow[dd, "\lambda b.\lambda c.(b{,}c)"] \\
& & \underset{\sharp}{\overset{\flat}{\substack{\longmapsfrom\\ \longmapsto}}} & & \\
D & D \arrow[rr, mapsto] \arrow[dd, "k"'] & & (C{\to}D) \arrow[dd, "\lambda f.\lambda c.k(fc)"] & (C{\to}(B{\times}C)) \\
& & \longmapsto & & \\
& E \arrow[rr, mapsto] & & (C{\to}E) &
\end{tikzcd}$$

$$\begin{tikzcd}
\mathbf{Set} \arrow[r, shift right, "(C\to)"'] & \mathbf{Set} \arrow[l, shift right, "(\times C)"']
\end{tikzcd}$$

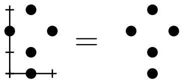
Figure 1: Three cases of internal views and external views.

was mostly based on exercises that the students would try to solve together by discussing on the whiteboard; it was mostly attended by Computer Science students who had just finished a course on Discrete Mathematics, but there were also some Psychology and Art students — that unfortunately left after the first weeks of each semester. All these students, including the CompSci ones, had in common that definitions only made sense to them after they had played with a few concrete examples; at some parts of the course I would ask them to read some sections of this paper, then work on some extra exercises that I had prepared, and then read excerpts of books like [Dal08] or [Awo06]. Most sections of this paper had been tested “on real children” in this way, and were rewritten several times after their feedback and reactions. I owe them many thanks — I'm glad that they had fun in the process – and I hope that I'll be able in the future to transform what I learned with them into precise techniques for writing “for children”.

1 Positional notations

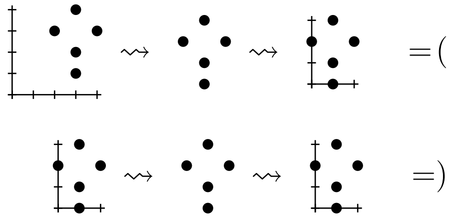
Definition: a *ZSet* is a finite, non-empty subset of $\mathbb{N}^2$ that touches both axes, i.e., that has a point of the form $(0, \_)$ and a point of the form $(\_, 0)$. We will often represent ZSets using a bullet notation, with or without the axes and ticks. For example:

$$K = \left\{ \begin{matrix} & \text{(1,3),} & \\ \text{(0,2),} & & \text{(2,2),} \\ & \text{(1,1),} & \\ & \text{(1,0)} & \end{matrix} \right\} =$$



We will use the ZSet above a lot in examples, so let's give it a short name: K (“kite”).

The condition of touching both axes is what lets us represent ZSets unambiguously using just the bullets:



We can use a positional notation to represent functions *from* a ZSet. For example, if

$$\begin{array}{rrcl} f & : & K & \to & \mathbb{N} \\ & & (x,y) & \mapsto & x \end{array}$$

then

$$f = \left\{ \begin{matrix} & \text{((1,3),1),} & \\ \text{((0,2),0),} & & \text{((2,2),2),} \\ & \text{((1,1),1),} & \\ & \text{((1,0),1)} & \end{matrix} \right\} = \begin{smallmatrix} & 1 & \\ 0 & & 2 \\ & 1 & \\ & 1 & \end{smallmatrix}$$

We will sometimes use λ-notation to represent functions compactly. For example:

$$\lambda(x,y){:}K.x = \left\{ \begin{smallmatrix} & ((1,3),1), & \\ ((0,2),0), & & ((2,2),2), \\ & ((1,1),1), & \\ & ((1,0),1) & \end{smallmatrix} \right\} = \begin{smallmatrix} & 1 & \\ 0 & & 2 \\ & 1 & \\ & 1 & \end{smallmatrix}$$

$$\lambda(x,y){:}K.y = \left\{ \begin{smallmatrix} & ((1,3),3), & \\ ((0,2),2), & & ((2,2),2), \\ & ((1,1),1), & \\ & ((1,0),0) & \end{smallmatrix} \right\} = \begin{smallmatrix} & 3 & \\ 2 & & 2 \\ & 1 & \\ & 0 & \end{smallmatrix}$$

The "reading order" on the points of a ZSet S "lists" the points of S starting from the top and going from left to right in each line. More precisely, if S has n points then $r_S : S \to \{1, \ldots, n\}$ is a bijection, and for example:

$$r_K = \begin{smallmatrix} & 1 & \\ 2 & & 3 \\ & 4 & \\ & 5 & \end{smallmatrix}$$

Subsets of a ZSet are represented with a notation with '$\bullet$'s and '$\cdot$', and partial functions from a ZSet are represented with '$\cdot$'s where they are not defined. For example:

$$\begin{smallmatrix} & \bullet & \\ \cdot & & \bullet \\ & \bullet & \\ & \cdot & \end{smallmatrix} \qquad \begin{smallmatrix} & 1 & \\ \cdot & & 3 \\ & 4 & \\ & \cdot & \end{smallmatrix}$$

The *characteristic function* of a subset S' of a ZSet S is the function $\chi_{S'} : S \to \{0,1\}$ that returns 1 exactly on the points of S'; for example, $\begin{smallmatrix} & 1 & \\ 0 & & 1 \\ & 1 & \\ & 0 & \end{smallmatrix}$ is the characteristic function of $\begin{smallmatrix} & \bullet & \\ \cdot & & \bullet \\ & \bullet & \\ & \cdot & \end{smallmatrix} \subset \begin{smallmatrix} & \bullet & \\ \bullet & & \bullet \\ & \bullet & \\ & \bullet & \end{smallmatrix}$. We will sometimes denote subsets by their characteristic functions because this makes them easier to "pronounce" by reading aloud their digits in the reading order — for example, $\begin{smallmatrix} & 1 & \\ 0 & & 1 \\ & 1 & \\ & 0 & \end{smallmatrix}$ is "one-zero-one-one-zero" (see sec.12).

2 ZDAGs

We will sometimes use the bullet notation for a ZSet S as a *shorthand* for one of the two DAGs induced by S: one with its arrows going up, the other one with them going down. For example: sometimes

$$\begin{tikzcd}
 & & & \bullet \arrow[dl] \arrow[dr] & & & & (1,3) \arrow[dl] \arrow[dr] & & \\
\begin{smallmatrix} & \bullet & \\ \bullet & & \bullet \\ & \bullet & \\ & \bullet & \end{smallmatrix} & \text{will stand for:} & \bullet \arrow[dr] & & \bullet \arrow[dl] & = & (0,2) \arrow[dr] & & (2,2) \arrow[dl] & = \\
 & & & \bullet \arrow[d] & & & & (1,1) \arrow[d] & & \\
 & & & \bullet & & & & (1,0) & &
\end{tikzcd}$$

$$\left(\left\{ \begin{smallmatrix} & (1,3), & \\ (0,2), & & (2,2), \\ & (1,1), & \\ & (0,0) & \end{smallmatrix} \right\}, \left\{ \begin{smallmatrix} ((1,3),(0,2)),((1,3),(2,2)), \\ ((0,2),(1,1)),((2,2),(1,1)), \\ ((1,1),(0,0)) \end{smallmatrix} \right\} \right)$$

Let's formalize this.

Consider a game in which black and white pawns are placed on points of $\mathbb{Z}^2$, and they can move like this:

Black pawns can move from (x, y) to $(x + k, y - 1)$ and white pawns from (x, y) to $(x + k, y + 1)$, where $k \in \{-1, 0, 1\}$. The mnemonic is that black pawns are "solid", and thus "heavy", and they "sink", so they move down; white pawns are "hollow", and thus "light", and they "float", so they move up.

Let's now restrict the board positions to a ZSet S. Black pawns can move from (x, y) to $(x + k, y - 1)$ and white pawns from (x, y) to $(x + k, y + 1)$, where $k \in \{-1, 0, 1\}$, but only when the starting and ending positions both belong to S. The sets of possible black pawn moves and white pawn moves on S can be defined formally as:

$$\begin{aligned}\mathsf{BPM}(S) &= \{\, ((x, y), (x', y')) \in S^2 \mid x - x' \in \{-1, 0, 1\}, y' = y - 1 \,\} \\ \mathsf{WPM}(S) &= \{\, ((x, y), (x', y')) \in S^2 \mid x - x' \in \{-1, 0, 1\}, y' = y + 1 \,\}\end{aligned}$$

...and now please forget everything else you expect from a game — like starting position, capturing, objective, winning... the idea of a "game" was just a tool to let us explain $\mathsf{BPM}(S)$ and $\mathsf{WPM}(S)$ quickly.

A *ZDAG* is a DAG of the form $(S, \mathsf{BPM}(S))$ or $(S, \mathsf{WPM}(S))$, where S is a ZSet.

A *ZPO* is partial order of the form $(S, \mathsf{BPM}(S)^*)$ or $(S, \mathsf{WPM}(S)^*)$, where S is a ZSet and the '*' denotes the transitive-reflexive closure of the relation.

Sometimes, when this is clear from the context, a bullet diagram like $\begin{smallmatrix} & \bullet & \\ \bullet & & \bullet \\ & \bullet & \\ & \bullet & \end{smallmatrix}$ will stand for either the ZDAGs $(\begin{smallmatrix} & \bullet & \\ \bullet & & \bullet \\ & \bullet & \\ & \bullet & \end{smallmatrix}, \mathsf{BPM}(\begin{smallmatrix} & \bullet & \\ \bullet & & \bullet \\ & \bullet & \\ & \bullet & \end{smallmatrix}))$ or $(\begin{smallmatrix} & \bullet & \\ \bullet & & \bullet \\ & \bullet & \\ & \bullet & \end{smallmatrix}, \mathsf{WPM}(\begin{smallmatrix} & \bullet & \\ \bullet & & \bullet \\ & \bullet & \\ & \bullet & \end{smallmatrix}))$, or for the ZPOs $(\begin{smallmatrix} & \bullet & \\ \bullet & & \bullet \\ & \bullet & \\ & \bullet & \end{smallmatrix}, \mathsf{BPM}(\begin{smallmatrix} & \bullet & \\ \bullet & & \bullet \\ & \bullet & \\ & \bullet & \end{smallmatrix})^*)$ or $(\begin{smallmatrix} & \bullet & \\ \bullet & & \bullet \\ & \bullet & \\ & \bullet & \end{smallmatrix}, \mathsf{WPM}(\begin{smallmatrix} & \bullet & \\ \bullet & & \bullet \\ & \bullet & \\ & \bullet & \end{smallmatrix})^*)$ (sec.4).

3 LR-coordinates

The *lr-coordinates* are useful for working on quarter-plane of $\mathbb{Z}^2$ that looks like $\mathbb{N}^2$ turned 45° to the left. Let $\langle l, r\rangle := (-l + r, l + r)$; then (the bottom part of) $\{\, \langle l, r\rangle \mid l, r \in \mathbb{N} \,\}$ is:

$$\begin{array}{ccccccccc} \langle 4,0\rangle && \langle 3,1\rangle && \langle 2,2\rangle && \langle 1,3\rangle && \langle 0,4\rangle \\ & \langle 3,0\rangle && \langle 2,1\rangle && \langle 1,2\rangle && \langle 0,3\rangle & \\ && \langle 2,0\rangle && \langle 1,1\rangle && \langle 0,2\rangle && \\ &&& \langle 1,0\rangle && \langle 0,1\rangle &&& \\ &&&& \langle 0,0\rangle &&&& \end{array} \quad = \quad \begin{array}{ccccccccc} (-4,4) && (-2,4) && (0,4) && (2,4) && (4,4) \\ & (-3,3) && (-1,3) && (1,3) && (3,3) & \\ && (-2,2) && (0,2) && (2,2) && \\ &&& (-1,1) && (1,1) &&& \\ &&&& (0,0) &&&& \end{array}$$

Sometimes we will write lr instead of $\langle l, r\rangle$. So:

$$\begin{array}{ccccccccc}
40 && 31 && 22 && 13 && 04 \\
& 30 && 21 && 12 && 03 & \\
&& 20 && 11 && 02 && \\
&&& 10 && 01 &&& \\
&&&& 00 &&&&
\end{array}
\quad = \quad
\begin{array}{ccccccccc}
(-4,4) && (-2,4) && (0,4) && (2,4) && (4,4) \\
& (-3,3) && (-1,3) && (1,3) && (3,3) & \\
&& (-2,2) && (0,2) && (2,2) && \\
&&& (-1,1) && (1,1) &&& \\
&&&& (0,0) &&&&
\end{array}$$

Let $\mathbb{LR} = \{\, \langle l, r\rangle \mid l, r \in \mathbb{N} \,\}$.

4 ZHAs

A *ZHA* is a subset of $\mathbb{LR}$ “between a left and a right wall”, as we will see.

A triple (h, L, R) is a “height-left-right-wall” when:
1) $h \in \mathbb{N}$
2) $L : \{0, \ldots, h\} \to \mathbb{Z}$ and $R : \{0, \ldots, h\} \to \mathbb{Z}$
3) $L(h) = R(h)$ (the top points of the walls are the same)
4) $L(0) = R(0) = 0$ (the bottom points of the walls are the same, 0)
5) $\forall y \in \{0, \ldots, h\}. L(y) \leq R(y)$ (“left” is left of “right”)
6) $\forall y \in \{1, \ldots, h\}. L(y) - L(y-1) = \pm 1$ (the left wall makes no jumps)
7) $\forall y \in \{1, \ldots, h\}. R(y) - R(y-1) = \pm 1$ (the right wall makes no jumps)

The *ZHA generated by* a height-left-right-wall (h, L, R) is the set of all points of $\mathbb{LR}$ with valid height and between the left and the right walls. Formally:

$$\mathsf{ZHAG}(h, L, R) = \{\, (x, y) \in \mathbb{LR} \mid y \leq h, L(y) \leq x \leq R(y) \,\}.$$

A *ZHA* is a set of the form $\mathsf{ZHAG}(h, L, R)$, where the triple (h, L, R) is a height-left-right-wall.

Here is an example of a ZHA (with the white pawn moves on it):

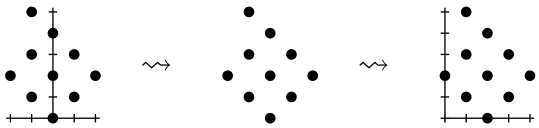
$$\begin{tikzcd}
 & (-3,9) & & & & \\
(-4,8) \arrow[ur] & & (-2,8) \arrow[ul] & & & \\
 & (-3,7) \arrow[ul] \arrow[ur] & & & & \\
 & & (-2,6) \arrow[ul] & & & \\
 & & & (-1,5) \arrow[ul] & & \\
 & & (-2,4) \arrow[ur] & & (0,4) \arrow[ul] & \\
 & (-3,3) \arrow[ur] & & (-1,3) \arrow[ul] \arrow[ur] & & (1,3) \arrow[ul] \\
 & & (-2,2) \arrow[ul] \arrow[ur] & & (0,2) \arrow[ul] \arrow[ur] & \\
 & & & (-1,1) \arrow[ul] \arrow[ur] & & (1,1) \arrow[ul] \\
 & & & & (0,0) \arrow[ul] \arrow[ur] &
\end{tikzcd}$$

$$L(9) = -3 \quad R(9) = -3 \quad L(9) = R(9) \quad h = 9$$
$$L(8) = -4 \quad R(8) = -2$$
$$L(7) = -3 \quad R(7) = -3$$
$$L(6) = -2 \quad R(6) = -2$$
$$L(5) = -1 \quad R(5) = -1$$
$$L(4) = -2 \quad R(4) = 0$$
$$L(3) = -3 \quad R(3) = 1$$
$$L(2) = -2 \quad R(2) = 0$$
$$L(1) = -1 \quad R(1) = 1$$
$$L(0) = 0 \quad R(0) = 0 \quad L(0) = R(0) = 0$$

We will see later (in section 7) that ZHAs (with white pawn moves) are Heyting Algebras.

5 Conventions on diagrams without axes

We can use a bullet notation to denote ZHAs, but look at what happens when we start with a ZHA, erase the axes, and then add the axes back using the convention from sec.1:



The new, restored axes are in a different position — the bottom point of the original ZHA at the left was $(0,0)$, but in the ZSet at the right the bottom point is $(2,0)$.

The convention from sec.1 is not adequate for ZHAs.

Let's modify it!

From this point on, the convention on where to draw the axes will be this one: *when it is clear from the context that a bullet diagram represents a ZHA,* then its (unique) bottom point has coordinate $(0,0)$, and we use that to draw the axes; otherwise we apply the old convention, that chooses $(0,0)$ as the point that makes the diagram fit in $\mathbb{N}^2$ and touch both axes.

The new convention with two cases also applies to functions from ZHAs, and to partial functions and subsets. For example:

$$B = \begin{array}{ccccc} & \bullet & & & \\ & & \bullet & & \\ & \bullet & & \bullet & \\ \bullet & & \bullet & & \bullet \\ & \bullet & & \bullet & \\ & & \bullet & & \end{array} \quad \text{(a ZHA)} \qquad \lambda(x,y){:}B.x = \begin{array}{ccccc} & \text{-1} & & & \\ & & 0 & & \\ & \text{-1} & & 1 & \\ \text{-2} & & 0 & & 2 \\ & \text{-1} & & 1 & \\ & & 0 & & \end{array}$$

$$\lambda\langle l,r\rangle{:}B.l = \begin{array}{ccccc} & 3 & & & \\ & & 2 & & \\ & 2 & & 1 & \\ 2 & & 1 & & 0 \\ & 1 & & 0 & \\ & & 0 & & \end{array} \qquad \lambda\langle l,r\rangle{:}B.r = \begin{array}{ccccc} & 2 & & & \\ & & 2 & & \\ & 1 & & 2 & \\ 0 & & 1 & & 2 \\ & 0 & & 1 & \\ & & 0 & & \end{array}$$

We will often denote ZHAs by the identity function on them:

$$\lambda\langle l,r\rangle{:}B.\langle l,r\rangle \;=\; \lambda lr{:}B.lr \;=\; \begin{array}{ccccc} & 32 & & & \\ & & 22 & & \\ & 21 & & 12 & \\ 20 & & 11 & & 02 \\ & 10 & & 01 & \\ & & 00 & & \end{array} \qquad B \;=\; \begin{array}{ccccc} & 32 & & & \\ & & 22 & & \\ & 21 & & 12 & \\ 20 & & 11 & & 02 \\ & 10 & & 01 & \\ & & 00 & & \end{array}$$

Note that we are using the compact notation from the end of section 3: 'lr' instead of '$\langle l,r\rangle$'.

6 Propositional calculus

A *PC-structure* is a tuple

$$L = (\Omega, \leq, \top, \bot, \land, \lor, \to, \leftrightarrow, \neg),$$

where:

Ω is the "set of truth values",
$\leq$ is a relation on Ω,
$\top$ and $\bot$ are two elements of Ω,
$\land$, $\lor$, $\to$, $\leftrightarrow$ are functions from $\Omega \times \Omega$ to Ω,
$\neg$ is a function from Ω to Ω.

Classical Logic "is" a PC-structure, with $\Omega = \{0,1\}$, $\top = 1$, $\bot = 0$, $\leq = \{(0,0),(0,1),(1,0)\}$, $\land = \left\{ \begin{smallmatrix} ((0,0),0),((0,1),0), \\ ((1,0),0),((1,1),1) \end{smallmatrix} \right\}$, etc.

PC-structures let us interpret expressions from Propositional Calculus ("PC-expressions"), and let us define a notion of *tautology*. For example, in Classical Logic,

- $\neg\neg P \leftrightarrow P$ is a tautology because it is valid (i.e., it yields $\top$) for all values of P in Ω,

- $\neg(P \land Q) \to (\neg P \lor \neg Q)$ is a tautology because it is valid for all values of P and Q in Ω,

- but $P \lor Q \to P \land Q$ is *not* a tautology, because when $P = 0$ and $Q = 1$ the result is not $\top$:

$$\underbrace{\underbrace{\underbrace{P}_{0} \lor \underbrace{Q}_{1}}_{1} \to \underbrace{\underbrace{P}_{0} \land \underbrace{Q}_{1}}_{0}}_{0}$$

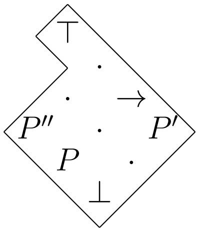
7 Propositional calculus in a ZHA

Let Ω be the set of points of a ZHA and $\leq$ the default partial order on it. The *default meanings for* $\top, \bot, \land, \lor, \to, \leftrightarrow, \neg$ are these ones:

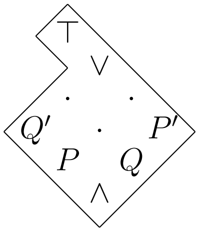
$$\begin{array}{rcl}
\langle a,b\rangle \leq \langle c,d\rangle & := & a \leq c \land b \leq d \\
\langle a,b\rangle \geq \langle c,d\rangle & := & a \geq c \land b \geq d \\
\\
\langle a,b\rangle \ \mathsf{above}\ \langle c,d\rangle & := & a \geq c \land b \geq d \\
\langle a,b\rangle \ \mathsf{below}\ \langle c,d\rangle & := & a \leq c \land b \leq d \\
\langle a,b\rangle \ \mathsf{leftof}\ \langle c,d\rangle & := & a \geq c \land b \leq d \\
\langle a,b\rangle \ \mathsf{rightof}\ \langle c,d\rangle & := & a \leq c \land b \geq d \\
\\
\mathsf{valid}(\langle a,b\rangle) & := & \langle a,b\rangle \in \Omega \\
\mathsf{ne}(\langle a,b\rangle) & := & \mathsf{if\ valid}\ (\langle a,b+1\rangle)\ \mathsf{then\ ne}(\langle a,b+1\rangle)\ \mathsf{else}\ \langle a,b\rangle\ \mathsf{end} \\
\mathsf{nw}(\langle a,b\rangle) & := & \mathsf{if\ valid}\ (\langle a+1,b\rangle)\ \mathsf{then\ nw}(\langle a+1,b\rangle)\ \mathsf{else}\ \langle a,b\rangle\ \mathsf{end} \\
\\
\langle a,b\rangle \land \langle c,d\rangle & := & \langle \mathsf{min}(a,c), \mathsf{min}(b,d)\rangle \\
\langle a,b\rangle \lor \langle c,d\rangle & := & \langle \mathsf{max}(a,c), \mathsf{max}(b,d)\rangle \\
\\
\langle a,b\rangle \to \langle c,d\rangle & := & \begin{array}{llll}
\mathsf{if} & \langle a,b\rangle\ \mathsf{below}\ \langle c,d\rangle & \mathsf{then} & \top \\
\mathsf{elseif} & \langle a,b\rangle\ \mathsf{leftof}\ \langle c,d\rangle & \mathsf{then} & \mathsf{ne}(\langle c,d\rangle) \\
\mathsf{elseif} & \langle a,b\rangle\ \mathsf{rightof}\ \langle c,d\rangle & \mathsf{then} & \mathsf{nw}(\langle c,d\rangle) \\
\mathsf{elseif} & \langle a,b\rangle\ \mathsf{above}\ \langle c,d\rangle & \mathsf{then} & \langle c,d\rangle \\
\mathsf{end} & & &
\end{array} \\
\top & := & \mathsf{sup}(\Omega) \\
\bot & := & \langle 0,0\rangle \\
\neg\langle a,b\rangle & := & \langle a,b\rangle \to \bot \\
\langle a,b\rangle \leftrightarrow \langle c,d\rangle & := & (\langle a,b\rangle \to \langle c,d\rangle) \land (\langle c,d\rangle \to \langle a,b\rangle)
\end{array}$$

Let Ω be the ZHA at the top left in the figure below. Then, with the default meanings for the connectives neither $\neg\neg P \to P$ nor $\neg(P \land Q) \to (\neg P \lor \neg Q)$ are tautologies, as there are valuations that make them yield results different than $\top = 32$:

$$\begin{array}{ccccc} & 32 & & & \\ & & 22 & & \\ & 21 & & 12 & \\ 20 & & 11 & & 02 \\ & 10 & & 01 & \\ & & 00 & & \end{array}$$



$$\underbrace{(\underbrace{\neg\underbrace{\neg\underbrace{P}_{10}}_{02}}_{20}) \to \underbrace{P}_{10}}_{12}$$



$$\underbrace{\underbrace{\neg(\underbrace{\underbrace{P}_{10}\land\underbrace{Q}_{01}}_{00})}_{32} \to \underbrace{(\underbrace{\neg\underbrace{P}_{10}}_{02}\lor\underbrace{\neg\underbrace{Q}_{01}}_{20})}_{22}}_{22}$$

So: *some* classical tautologies are not tautologies in this ZHA.

The somewhat arbitrary-looking definition of '→' will be explained at the end of the next section.

8 Heyting Algebras

A *Heyting Algebra* is a PC-structure

$$H = (\Omega, \leq_H, \top_H, \bot_H, \land_H, \lor_H, \to_H, \leftrightarrow_H, \neg_H),$$

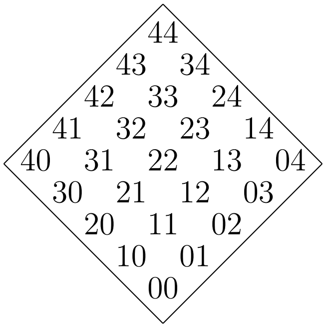
in which:

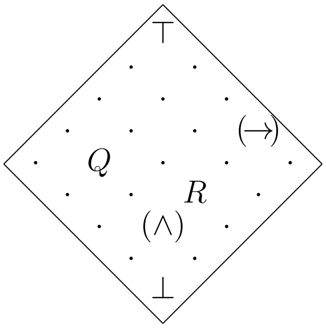
1) $(\Omega, \leq_H)$ is a partial order
2) $\top_H$ is the top element of the partial order
3) $\bot_H$ is the bottom element of the partial order
4) $P \leftrightarrow_H Q$ is the same as $(P \to_H Q) \land_H (Q \to_H P)$
5) $\neg_H P$ is the same as $P \to_H \bot_H$
6) $\forall P, Q, R \in \Omega.\ (P \leq_H (Q \land_H R)) \leftrightarrow ((P \leq_H Q) \land (P \leq_H R))$
7) $\forall P, Q, R \in \Omega.\ ((P \lor_H Q) \leq_H R) \leftrightarrow ((P \leq_H R) \land (Q \leq_H R))$
8) $\forall P, Q, R \in \Omega.\ (P \leq_H (Q \to_H R)) \leftrightarrow ((P \land_H Q) \leq_H R)$
6') $\forall Q, R \in \Omega. \exists! Y \in \Omega. \forall P \in \Omega.\ (P \leq_H Y) \leftrightarrow ((P \leq_H Q) \land (P \leq_H R))$
7') $\forall P, Q \in \Omega. \exists! X \in \Omega. \forall R \in \Omega.\ (X \leq_H R) \leftrightarrow ((P \leq_H R) \land (Q \leq_H R))$
8') $\forall Q, R \in \Omega. \exists! Y \in \Omega. \forall P \in \Omega.\ (P \leq_H Y) \leftrightarrow ((P \land_H R) \leq_H R)$

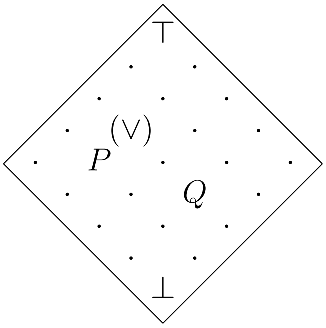
The conditions 6', 7', 8' say that there are unique elements in Ω that "behave as" $Q \land_H R$, $P \lor_H Q$ and $Q \to_H R$ for given P, Q, R; the conditions 6,7,8 say that $Q \land_H R$, $P \lor_H Q$ and $Q \to_H R$ are exactly the elements with this behavior.

The positional notation on ZHAs is very helpful for visualizing what the conditions 6',7',8',6,7,8 "mean". More precisely: once we fix a ZHA Ω and truth-values $P, Q, R \in \Omega$ we have a way to draw and to visualize the "behavior" of each subexpression of the conditions 6, 7, 8 using the positional notations of sec.1, and we can use that to obtain the only possible values for $Q \land_H R$, $P \lor_H Q$ and $Q \to_H R$.

We will examine three particular cases: with Ω being the ZDAG on the left below,







a) if $Q = 31$ and $R = 12$ then $Q \land_H R = 11$,
b) if $P = 31$ and $Q = 12$ then $P \lor_H Q = 32$,
c) if $Q = 31$ and $R = 12$ then $Q \to_H R = 14$.

Before we start, note that in 6, 7, 8, 6', 7', 8' some subexpressions yield truth values in Ω and other subexpressions yield standard truth values. For example, in 6, with $P = 20$, we have:

$$\underbrace{(\underbrace{\underbrace{P}_{20} \le_H (\underbrace{\underbrace{Q}_{31} \land_H \underbrace{R}_{12}}_{11}))}_{0} \leftrightarrow \underbrace{((\underbrace{\underbrace{P}_{20} \le_H \underbrace{Q}_{31}}_{1}) \land (\underbrace{\underbrace{P}_{20} \le_H \underbrace{R}_{12}}_{0}))}_{0}}_{1}$$

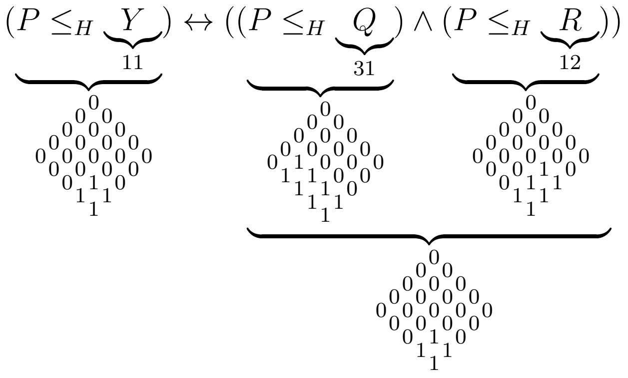
Case (a). Let $Q = 31$ and $R = 12$. We want to see that $Q \land_H R = 11$, i.e., that

$$\forall P \in \Omega.\ (P \le_H Y) \leftrightarrow ((P \le_H Q) \land (P \le_H R))$$

holds for $Y = 11$ and for no other $Y \in \Omega$. We can visualize the behavior of $P \le_H Q$ for all 'P's by drawing $\lambda P{:}\Omega.(P \le_H Q)$ in the positional notation; then we do the same for $\lambda P{:}\Omega.(P \le_H R)$ and for $\lambda P{:}\Omega.((P \le_H Q) \land (P \le_H R))$. Suppose that the full expression, '$\forall P{:}\Omega.\underline{\quad\quad}$', is true; then the behavior of the left side of the '$\leftrightarrow$', $\lambda P{:}\Omega.(P \le_H Y)$, has to be a copy of the behavior of the right side, and that lets us find the only adequate value for Y.

The order in which we calculate and draw things is below, followed by the results themselves:

$$\underbrace{(P \leq_H \underbrace{Y}_{(7)})}_{(6)} \leftrightarrow (\underbrace{\underbrace{(P \leq_H \underbrace{Q}_{(1)})}_{(3)} \land \underbrace{(P \leq_H \underbrace{R}_{(2)})}_{(4)}}_{(5)})$$

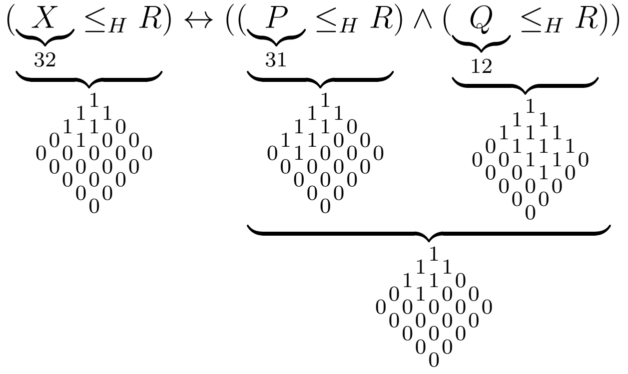


Case (b). Let $P = 31$ and $Q = 12$. We want to see that $P \vee_H Q = 32$, i.e., that

$$\forall R{:}\Omega.\ \ (X \leq_H R) \leftrightarrow ((P \leq_H R) \land (Q \leq_H R))$$

holds for $X = 32$ and for no other $X \in \Omega$. We do essentially the same as we did in (a), but now we calculate $\lambda R{:}\Omega.(P \leq_H R)$, $\lambda R{:}\Omega.(Q \leq_H R)$, and $\lambda R{:}\Omega.((P \leq_H R) \land (Q \leq_H R))$. The order in which we calculate and draw things is below, followed by the results themselves:

$$\underbrace{(\underbrace{X}_{(7)} \leq_H R)}_{(6)} \leftrightarrow (\underbrace{\underbrace{(\underbrace{P}_{(1)} \leq_H R)}_{(3)} \land \underbrace{(\underbrace{Q}_{(2)} \leq_H R)}_{(4)}}_{(5)})$$

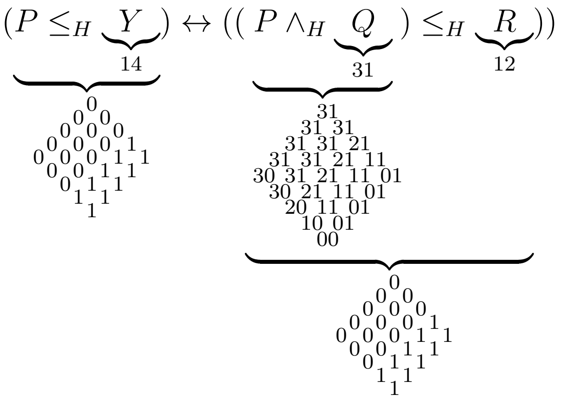


Case (c). Let $Q = 31$ and $R = 12$. We want to see that $Q \to_H R = 14$, i.e., that

$$\forall P{:}\Omega.\ \ (P \leq_H Y) \leftrightarrow ((P \wedge_H Q) \leq_H R)$$

holds for $Y = 14$ and for no other $Y \in \Omega$. Here we have to do something slightly different. We start by visualizing $\lambda P{:}\Omega.(P \wedge_H Q)$, which is a function from Ω to Ω, not a function from Ω to $\{0,1\}$ like the ones we were using before. The order in which we calculate and draw things is below, followed by the results:

$$\underbrace{(P \leq_H \underbrace{Y}_{(6)})}_{(5)} \leftrightarrow (\underbrace{\underbrace{(P \wedge_H \underbrace{Q}_{(1)})}_{(3)} \leq_H \underbrace{R}_{(2)}}_{(4)})$$



9 The two implications are equivalent

In sec.7 we gave a definition of '$\to$' that is easy to calculate, and in sec.8 we saw a way to find by brute force[1] a value for $Q \to R$ that obeys

$$(P \leq (Q \to R)) \leftrightarrow (P \wedge Q \leq R)$$

for all P. In this section we will see a proof that these two operations — called '$\overset{\text{C}}{\to}$' and '$\overset{\text{HA}}{\to}$' from here on — always give the same results.

Theorem 9.1 *We have* $(Q \overset{C}{\to} R) = (Q \overset{HA}{\to} R)$, *for any ZHA* H *and* $Q, R \in H$.

---

[1] "When in doubt use brute force" — Ken Thompson

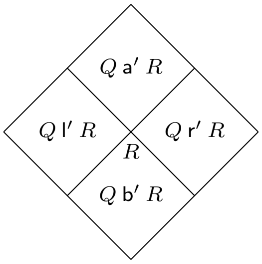
The proof will take the rest of this section, and our approach will be to check that for any ZHA H and $Q, R \in H$ this holds, for all $P \in H$:

$$(P \leq (Q \overset{\mathrm{C}}{\to} R)) \leftrightarrow (P \wedge Q \leq R).$$

In '$\overset{\mathrm{C}}{\to}$' the order of the cases is very important. For example, if $cd = 21$ and $ef = 23$ then both "$cd\,\mathsf{below}\,ef$" and "$cd\,\mathsf{leftof}\,ef$" are true, but "$cd\,\mathsf{below}\,ef$" takes precedence and so $cd \overset{\mathrm{C}}{\to} ef = \top$. We can fix this by creating variants of below, leftof, righof and above, called below', leftof', righof' and above', that make the four cases disjoint. Abbreviating below, leftof, righof and above as b, l, r and a, we have:

$$\begin{array}{rcl@{\qquad}rcl}
cd\ \mathsf{b}\ ef & := & c \leq e \wedge d \leq f & cd\ \mathsf{b}'\ ef & := & c \leq e \wedge d \leq f \\
cd\ \mathsf{l}\ ef & := & c \leq e \wedge d \geq f & cd\ \mathsf{l}'\ ef & := & c \leq e \wedge d > f \\
cd\ \mathsf{r}\ ef & := & c \geq e \wedge d \leq f & cd\ \mathsf{r}'\ ef & := & c > e \wedge d \leq f \\
cd\ \mathsf{a}\ ef & := & c > e \wedge d > f & cd\ \mathsf{a}'\ ef & := & c > e \wedge d > f
\end{array}$$

visually the regions are these, for R fixed:



Note that R belongs to the lower region — i.e., $R\,\mathsf{b}'\,R$.

Now we clearly have:

$$Q \overset{\mathrm{C}}{\to} R \ = \left(\begin{array}{llll}
\mathsf{if} & Q\,\mathsf{b}\,R & \mathsf{then} & \top \\
\mathsf{elseif} & Q\,\mathsf{l}\,R & \mathsf{then} & \mathsf{ne}(R) \\
\mathsf{elseif} & Q\,\mathsf{r}\,R & \mathsf{then} & \mathsf{nw}(R) \\
\mathsf{elseif} & Q\,\mathsf{a}\,R & \mathsf{then} & R \\
\mathsf{end} & & &
\end{array}\right) = \left(\begin{array}{llll}
\mathsf{if} & Q\,\mathsf{b}'\,R & \mathsf{then} & \top \\
\mathsf{elseif} & Q\,\mathsf{l}'\,R & \mathsf{then} & \mathsf{ne}(R) \\
\mathsf{elseif} & Q\,\mathsf{r}'\,R & \mathsf{then} & \mathsf{nw}(R) \\
\mathsf{elseif} & Q\,\mathsf{a}'\,R & \mathsf{then} & R \\
\mathsf{end} & & &
\end{array}\right)$$

and $P \leq (Q \overset{\mathrm{C}}{\to} R)$ can be expressed as a conjunction of the four cases:

$$
\begin{array}{l}
((P \leq Q \overset{\mathrm{C}}{\to} R) \leftrightarrow (P \land Q \leq R)) \\
\leftrightarrow \quad \left(\begin{array}{ll}
Q \, \mathsf{b}' \, R \to ((P \leq Q \overset{\mathrm{C}}{\to} R) \leftrightarrow (P \land Q \leq R)) & \land \\
Q \, \mathsf{l}' \, R \to ((P \leq Q \overset{\mathrm{C}}{\to} R) \leftrightarrow (P \land Q \leq R)) & \land \\
Q \, \mathsf{r}' \, R \to ((P \leq Q \overset{\mathrm{C}}{\to} R) \leftrightarrow (P \land Q \leq R)) & \land \\
Q \, \mathsf{a}' \, R \to ((P \leq Q \overset{\mathrm{C}}{\to} R) \leftrightarrow (P \land Q \leq R)) &
\end{array} \right) \\
\\
\leftrightarrow \quad \left(\begin{array}{ll}
Q \, \mathsf{b}' \, R \to ((P \leq \top) \leftrightarrow (P \land Q \leq R)) & \land \\
Q \, \mathsf{l}' \, R \to ((P \leq \mathsf{ne}(R)) \leftrightarrow (P \land Q \leq R)) & \land \\
Q \, \mathsf{r}' \, R \to ((P \leq \mathsf{nw}(R)) \leftrightarrow (P \land Q \leq R)) & \land \\
Q \, \mathsf{a}' \, R \to ((P \leq R) \leftrightarrow (P \land Q \leq R)) &
\end{array} \right)
\end{array}
$$

Let's introduce a notation: a "$\widehat{a}$" means "make this digit as big possible without leaving the ZHA". So,

$$
\text{in} \quad
\begin{array}{cccccc}
 & 54 & & & & \\
53 & & 44 & & & \\
 & 43 & & 34 & & \\
42 & & 33 & & 24 & \\
 & 32 & & 23 & & \\
31 & & 22 & & 13 & \\
 & 21 & & 12 & & 03 \\
20 & & 11 & & 02 & \\
 & 10 & & 01 & & \\
 & & 00 & & &
\end{array}
\quad \text{we have} \quad
\begin{array}{ccccc}
\widehat{1}\widehat{2} & = & 54 & = & \top, \\
1\widehat{2} & = & 13 & = & \mathsf{ne}(12), \\
\widehat{1}2 & = & 42 & = & \mathsf{nw}(12);
\end{array}
$$

This lets us rewrite $\top$ as $\widehat{e}\widehat{f}$, $\mathsf{ne}(ef)$ as $e\widehat{f}$, and $\mathsf{nw}(ef)$ as $\widehat{e}f$.

Making $P = ab$, $Q = cd$, $R = ef$, we have:

$$
\begin{array}{ll}
\multicolumn{2}{l}{((ab \leq cd \overset{\mathrm{C}}{\to} ef) \leftrightarrow (ab \land cd \leq ef))} \\
\leftrightarrow & \left(\begin{array}{ll}
cd\,\mathsf{b}'\,ef \to ((ab \leq \widehat{e}\widehat{f}) \leftrightarrow (ab \land cd \leq ef)) & \land \\
cd\,\mathsf{l}'\,ef \to ((ab \leq e\widehat{f}) \leftrightarrow (ab \land cd \leq ef)) & \land \\
cd\,\mathsf{r}'\,ef \to ((ab \leq \widehat{e}f) \leftrightarrow (ab \land cd \leq ef)) & \land \\
cd\,\mathsf{a}'\,ef \to ((ab \leq ef) \leftrightarrow (ab \land cd \leq ef)) &
\end{array} \right) \\
\\
\leftrightarrow & \left(\begin{array}{ll}
c \leq e \land d \leq f \to ((ab \leq \widehat{e}\widehat{f}) \leftrightarrow (ab \land cd \leq ef)) & \land \\
c > e \land d \leq f \to ((ab \leq e\widehat{f}) \leftrightarrow (ab \land cd \leq ef)) & \land \\
c \leq e \land d > f \to ((ab \leq \widehat{e}f) \leftrightarrow (ab \land cd \leq ef)) & \land \\
c > e \land d > f \to ((ab \leq ef) \leftrightarrow (ab \land cd \leq ef)) &
\end{array} \right) \\
\\
\leftrightarrow & \left(\begin{array}{ll}
c \leq e \land d \leq f \to ((ab \leq \widehat{e}\widehat{f}) \leftrightarrow (ab \land cd \leq cd)) & \land \\
c > e \land d \leq f \to ((ab \leq e\widehat{f}) \leftrightarrow (ab \land cd \leq ed)) & \land \\
c \leq e \land d > f \to ((ab \leq \widehat{e}f) \leftrightarrow (ab \land cd \leq cf)) & \land \\
c > e \land d > f \to ((ab \leq ef) \leftrightarrow (ab \land cd \leq ef)) &
\end{array} \right) \\
\\
\leftrightarrow & \left(\begin{array}{ll}
c \leq e \land d \leq f \to ((ab \leq \widehat{e}\widehat{f}) \leftrightarrow \top) & \land \\
c > e \land d \leq f \to ((ab \leq e\widehat{f}) \leftrightarrow a \leq e) & \land \\
c \leq e \land d > f \to ((ab \leq \widehat{e}f) \leftrightarrow b \leq f) & \land \\
c > e \land d > f \to ((ab \leq ef) \leftrightarrow (a \leq e \land b \leq f)) &
\end{array} \right)
\end{array}
$$

In the last conjunction the four cases are trivial to check.

10 Logic in a Heyting Algebra

In sec.8 we saw a set of conditions — called 1 to 8' — that characterize the "Heyting-Algebra-ness" of a PC-structure. It is easy to see that Heyting-Algebra-ness, or "HAness", is equivalent to this set of conditions:

$$\boxed{\begin{array}{llrclcll}
1) & \forall P. & (P \le P) & & & & & (\mathrm{id}) \\
 & \forall P,Q,R. & (P \le R) & \leftarrow & (P \le Q) & \land & (Q \le R) & (\mathrm{comp}) \\
2) & \forall P. & (P \le \top) & & & & & (\top_1) \\
3) & \forall Q. & (\bot \le Q) & & & & & (\bot_1) \\
6) & \forall P,Q,R. & (P \le Q \land R) & \rightarrow & (P \le Q) & & & (\land_1) \\
 & \forall P,Q,R. & (P \le Q \land R) & \rightarrow & & & (P \le R) & (\land_2) \\
 & \forall P,Q,R. & (P \le Q \land R) & \leftarrow & (P \le Q) & \land & (P \le R) & (\land_3) \\
7) & \forall P,Q,R. & (P \lor Q \le R) & \rightarrow & (P \le R) & & & (\lor_1) \\
 & \forall P,Q,R. & (P \lor Q \le R) & \rightarrow & & & (Q \le R) & (\lor_2) \\
 & \forall P,Q,R. & (P \lor Q \le R) & \leftarrow & (P \le R) & \land & (Q \le R) & (\lor_3) \\
8) & \forall P,Q,R. & (P \le Q {\to} R) & \rightarrow & (P \land Q \le R) & & & (\to_1) \\
 & \forall P,Q,R. & (P \le Q {\to} R) & \leftarrow & (P \land Q \le R) & & & (\to_2)
\end{array}}$$

We omitted the conditions 4 and 5, that defined '↔' and '¬' in terms of the other operators. The last column of the table gives a name to each of these new conditions.

These new conditions let us put (some) proofs about HAs in tree form, as we shall see soon.

Let us introduce two new notations. The first one,

$$(\mathrm{expr}) \begin{bmatrix} \scriptstyle v_1 := \mathrm{repl}_1 \\ \scriptstyle v_2 := \mathrm{repl}_2 \end{bmatrix}$$

indicates simultaneous substitution of all (free) occurrences of the variables v_1 and v_2 in expr by the replacements repl_1 and repl_2. For example,

$$((x+y)\cdot z) \begin{bmatrix} \scriptstyle x:=a+y \\ \scriptstyle y:=b+z \\ \scriptstyle z:=c+x \end{bmatrix} \quad = \quad ((a+y)+(b+z))\cdot(c+x).$$

The second is a way to write '→'s as horizontal bars. In

$$\frac{A \quad B \quad C}{D}\,\alpha \qquad \frac{E \quad F}{G}\,\beta \qquad \frac{H}{I}\,\gamma \qquad \frac{}{J}\,\delta \qquad \frac{\dfrac{}{K}\,\epsilon \quad \dfrac{L \quad M}{N}\,\zeta \quad O}{P}\,\eta$$

the trees mean:

- if A, B, C are true then D is true (by α),
- if E, F, are true then G is true (by β),

- if H is true then I is true (by γ),
- J is true (by δ, with no hypotheses),
- K is true (by ϵ); if L and M then N (by ζ); if K, N, O, then P (by η); combining all this we get a way to prove that if L, M, O, then P,

where $\alpha, \beta, \gamma, \delta, \epsilon, \zeta, \eta$ are usually names of rules.

The implications in the table in the beginning of this section can be rewritten as "tree rules" as:

$$\dfrac{}{P \le P}\,\text{id} \qquad \dfrac{P \le Q \quad Q \le R}{P \le R}\,\text{comp} \qquad \dfrac{}{P \le \top}\,\top_1 \qquad \dfrac{}{\bot \le Q}\,\bot_1$$

$$\dfrac{P \le Q \land R}{P \le Q}\,\land_1 \qquad \dfrac{P \le Q \land R}{P \le R}\,\land_2 \qquad \dfrac{P \le Q \quad P \le R}{P \le Q \land R}\,\land_3$$

$$\dfrac{P \lor Q \le R}{P \le R}\,\lor_1 \qquad \dfrac{P \lor Q \le R}{Q \le R}\,\lor_2 \qquad \dfrac{P \le R \quad Q \le R}{P \lor Q \le R}\,\lor_3$$

$$\dfrac{P \le Q {\to} R}{P \land Q \le R}\,{\to}_1 \qquad \dfrac{P \land Q \le R}{P \le Q {\to} R}\,{\to}_2$$

Note that the '$\forall P, Q, R \in \Omega$'s are left implicit in the tree rules, which means that *every substitution instance of the tree rules hold*; sometimes — but rarely — we will indicate the substitution explicitly, like this,

$$\left(\dfrac{P \land Q \le R}{P \le Q {\to} R}\,{\to}_2\right)\left[\begin{smallmatrix} Q:=P\to\bot \\ R:=\bot \end{smallmatrix}\right] \quad \rightsquigarrow \quad \dfrac{P \land (P{\to}\bot) \le \bot}{P \le ((P{\to}\bot){\to}\bot)}\,{\to}_2$$

$$({\to}_2)\left[\begin{smallmatrix} Q:=P\to\bot \\ R:=\bot \end{smallmatrix}\right] \quad \rightsquigarrow \quad \dfrac{P \land (P{\to}\bot) \le \bot}{P \le ((P{\to}\bot){\to}\bot)}\,{\to}_2\left[\begin{smallmatrix} Q:=P\to\bot \\ R:=\bot \end{smallmatrix}\right]$$

Usually we will only say '$\to_2$' instead of '$\to_2 \left[\begin{smallmatrix} Q:=P\to\bot \\ R:=\bot \end{smallmatrix}\right]$' at the right of a bar, and the task of discovering which substitution has been used is left to the reader.

The tree rules can be composed in a nice visual way. For example, this tree — let's call it $(\land\land)$,

$$\dfrac{\dfrac{\dfrac{\dfrac{}{P \land Q \le P \land Q}\,\text{id}}{P \land Q \le P}\,\land_1 \quad P \le R}{P \land Q \le R}\,\text{comp} \qquad \dfrac{\dfrac{\dfrac{}{P \land Q \le P \land Q}\,\text{id}}{P \land Q \le Q}\,\land_2 \quad Q \le S}{P \land Q \le S}\,\text{comp}}{P \land Q \le R \land S}\,\land_3$$

“is” a proof for:

$$\forall P, Q, R, S \in \Omega.\ (P \le R) \wedge (Q \le S) \to ((P \wedge Q) \le (R \wedge S)).$$

We can perform substitutions on trees, and the notation will be the same as for tree rules: for example, $(\wedge\wedge)\left[\,{\scriptstyle S:=P\wedge Q}\,\right]$.

10.1 Derived rules

Let be HAT the set of “Heyting Algebra rules in tree form” from the last section:

$$\mathsf{HAT} = \{(\mathrm{id}), \dots, (\to_2)\}.$$

Let’s see a way to treat HAT as a deductive system.
If S is a set of tree rules, we will write:

- $\mathsf{Trees}(\mathsf{S})$ for the set of all trees whose bars are all substituion instances of rules in S,
- $\mathsf{Trees}(\mathsf{S}, \{H_1, \dots, H_n\})$ for the set of all trees in $\mathsf{Trees}(\mathsf{S})$ whose hypotheses are contained in the set $\{H_1, \dots, H_n\}$, and
- $\mathsf{Trees}(\mathsf{S}, \{H_1, \dots, H_n\}, C)$ for the set of trees in $\mathsf{Trees}(\mathsf{S}, \{H_1, \dots, H_n\})$ having C as their conclusion.

When the set S is clear from the context, we write

$$\begin{array}{c} H_1 \quad \dots \quad H_n \\ \hline\hline C \end{array}$$

to mean: *we know* a tree in $\mathsf{Trees}(\mathsf{S}, \{H_1, \dots, H_n\}, C)$, and this is an abbreviation for it. I like to think of the double bar as the bellows of a closed accordion: when the accordion is closed we can still see the keyboards at both sides, but not the drawings painted on the folded part of the pleated cloth.

The notation that defines a *derived rule* is “*newrule* := *expansion*”, where *expansion* is a tree in $\mathsf{Trees}(\mathsf{S}, \{H_1, \dots, H_n\}, C)$ and *newrule* is a bar with hypotheses $H_1, \dots, H_n$ and conclusion C, written with a single bar with a (new) rule name, instead of with a double bar. For example, this is a version of Modus Ponens for Heyting Algebras:

$$\dfrac{P \le Q \quad P \le Q{\to}R}{P \le R}\,\mathsf{MP} \qquad := \qquad \dfrac{\dfrac{P \le Q{\to}R \quad P \le Q}{P \le (Q{\to}R){\wedge}Q}\,\wedge_3 \qquad \dfrac{\dfrac{}{Q{\to}R \le Q{\to}R}\,\mathrm{id}}{(Q{\to}R){\wedge}Q \le R}\,{\to_1}}{P \le R}\,\mathrm{comp}$$

After the definition of a derived rule — say, "$D_1 := E_1$" — the set of allowed tree rules that is implicit from the context is increased, with D_1 being added to it; when we define another derived rule, $D_2 := E_2$, its expansion E_2 can have bars that are substitution instances of D_1. After adding more derived rules, $D_3 := E_3$, …, $D_n := E_n$, we can use all the new rules $D_1, \ldots, D_n$ in our trees — and we have a way to remove all the derived rules from our trees! Take a tree $T \in \mathsf{Trees}(\mathsf{S} \cup \{D_1, \ldots, D_n\})$; replace each substitution instance of D_n in it by its expansion, then replace every substitution instance of D_{n-1} in the new tree by its expansion, and so on; after replacing all substitution instances of D_1 we get a tree in $\mathsf{Trees}(\mathsf{S})$, with the same hypotheses and the same conclusion as the original T.

We want to add these other derived rules:

$$\dfrac{}{Q \land R \le Q}\,{\land}E_1 \quad := \quad \dfrac{\dfrac{}{Q \land R \le Q \land R}\,\text{id}}{Q \land R \le Q}\,\land_1$$

$$\dfrac{}{Q \land R \le R}\,{\land}E_2 \quad := \quad \dfrac{\dfrac{}{Q \land R \le Q \land R}\,\text{id}}{Q \land R \le R}\,\land_2$$

$$\dfrac{}{P \le P \lor Q}\,{\lor}I_1 \quad := \quad \dfrac{\dfrac{}{P \lor Q \le P \lor Q}\,\text{id}}{P \le P \lor Q}\,\lor_1$$

$$\dfrac{}{Q \le P \lor Q}\,{\lor}I_2 \quad := \quad \dfrac{\dfrac{}{P \lor Q \le P \lor Q}\,\text{id}}{Q \le P \lor Q}\,\lor_2$$

$$\dfrac{P \land R \le S \quad Q \land R \le S}{(P \lor Q) \land R \le R}\,{\lor}E \quad := \quad \dfrac{\dfrac{\dfrac{P \land R \le S}{P \le R \to S}\,{\to_2} \quad \dfrac{Q \land R \le S}{Q \le R \to S}\,{\to_2}}{P \lor Q \le R \to S}\,\lor_3}{(P \lor Q) \land R \le R}\,{\to_1}$$

10.2 Natural deduction

The system HAT with all the derived rules of the last section added to it will be called HAND:

$$\mathsf{HAND} = \{(\text{id}), \ldots, (\to_2), (\mathsf{MP}), \ldots, ({\lor}E))\}$$

Trees in Natural Deduction for IPL can be translated into HAND by a method that we will sketch below. Note that this section is not self-contained — it should be regarded as an introduction to [NP01]. Note that *all our trees can be intepreted as proofs about Heyting Algebras.*

This is an example of a tree in Natural Deduction:

$$\dfrac{\dfrac{\dfrac{[P]^1 \quad P{\to}Q}{Q}\;(\to E) \qquad \dfrac{[P]^1 \quad P{\to}R}{R}\;(\to E)}{Q{\land}R}\;(\land I)}{P{\to}(Q{\land}R)}\;(\to I);1$$

The ";1" in its last bar means: below this point the hypotheses marked with '$[\,\cdot\,]^1$' are "discharged" from the list of hypotheses. Each subtree of a ND tree with undischarged hypotheses $H_1, \ldots, H_n$ and conclusion C will be interpreted as *some* tree in HAND with no hypotheses and conclusion $H_1 \land \ldots \land H_n \le C$ — there are usually several possible choices. So:

$$\dfrac{P \quad P{\to}Q}{Q} \quad\Longrightarrow\quad \dfrac{}{P \land (P{\to}Q) \le Q}\;\mathsf{MP}$$

$$\dfrac{P \quad P{\to}R}{R} \quad\Longrightarrow\quad \dfrac{}{P \land (P{\to}R) \le R}\;\mathsf{MP}$$

$$\dfrac{Q \quad R}{Q{\land}R} \quad\Longrightarrow\quad \dfrac{}{Q{\land}R \le Q{\land}R}\;\mathrm{id}$$

$$\dfrac{\dfrac{P \quad P{\to}Q}{Q} \qquad \dfrac{P \quad P{\to}R}{R}}{Q{\land}R} \quad\Longrightarrow\quad \overline{\overline{((P{\to}R) \land (P{\to}Q)) \land P \le Q{\land}R}}$$

$$\dfrac{\dfrac{\dfrac{[P]^1 \quad P{\to}Q}{Q} \qquad \dfrac{[P]^1 \quad P{\to}R}{R}}{Q{\land}R}}{P{\to}(Q{\land}R)}\;(\to I);1 \quad\Longrightarrow\quad \dfrac{\overline{\overline{((P{\to}R) \land (P{\to}Q)) \land P \le Q{\land}R}}}{(P{\to}R) \land (P{\to}Q) \le P{\to}Q{\land}R}\;{\to_2}$$

The ND rules that are difficult to understand and difficult to translate are the ones that involve discharges: '$(\to I)$', that appears above, and '$(\lor E)$':

$$\dfrac{P \lor Q \qquad \begin{array}{c}[P]^1 \quad R \\ \vdots\; T_1 \\ S\end{array} \qquad \begin{array}{c}[Q]^1 \quad R \\ \vdots\; T_2 \\ S\end{array}}{S}\;(\lor E) \quad\Longrightarrow\quad \dfrac{\overline{\overline{P \land R \le S}}\;T_1 \qquad \overline{\overline{Q \land R \le S}}\;T_2}{(P \lor Q) \land R \le S}\;\lor E$$

Note that the derived rule $\lor E$ is used to combine the translations of the subtrees T_1 and T_2 into a translation of the whole ND tree.

My suggestion for the readers that are seeing this for the first time is: start by translating the ND tree below

$$\dfrac{\dfrac{(P\lor Q)\land R}{P\lor Q}\,(\land E_1) \qquad \dfrac{\dfrac{[P]^1 \quad \dfrac{(P\lor Q)\land R}{R}\,(\land E_2)}{P\land R}\,(\land I)}{(P\land R)\lor(Q\land R)}\,(\lor I_1) \qquad \dfrac{\dfrac{[Q]^1 \quad \dfrac{(P\lor Q)\land R}{R}\,(\land E_2)}{Q\land R}\,(\land I)}{(P\land R)\lor(Q\land R)}\,(\lor I_2)}{(P\land R)\lor(Q\land R)}\,(\lor E);1$$

to a tree in HAND, and then to a tree in HAT; then read the relevant parts of [NP01] to see how they would do that translation.

11 Topologies

The best way to connect ZHAs to several standard concepts is by seeing that ZHAs are topologies on certain finite sets — actually on 2-column acyclical graphs (sec.14). This will be done here and in the next few sections.

A *topology* on a set X is a subset $\mathcal{U}$ of $\mathcal{P}(X)$ that contains the "everything" and the "nothing" and is closed by binary unions and intersections and by arbitrary unions. Formally:

1) $\mathcal{U}$ contains X and $\varnothing$,

2) if $P,Q\in\mathcal{U}$ then $\mathcal{U}$ contains $P\cup Q$ and $P\cap Q$,

3) if $\mathcal{V}\subset\mathcal{U}$ then $\mathcal{U}$ contains $\bigcup\mathcal{V}$.

A *topological space* is a pair $(X,\mathcal{U})$ where X is a set and $\mathcal{U}$ is a topology on X.

When $(X,\mathcal{U})$ is a topological space and $U\in\mathcal{U}$ we say that U is *open* in $(X,\mathcal{U})$.

For example, let X be the ZSet $\begin{smallmatrix}\bullet&&\bullet\\&\bullet&\\\bullet&&\bullet\end{smallmatrix}$, and let's use the characteristic function notation from sec.1 to denote its subsets — we write $X=\begin{smallmatrix}1&&1\\&1&\\1&&1\end{smallmatrix}$ and $\varnothing=\begin{smallmatrix}0&&0\\&0&\\0&&0\end{smallmatrix}$ instead of $X=\begin{smallmatrix}\bullet&&\bullet\\&\bullet&\\\bullet&&\bullet\end{smallmatrix}$ and $\varnothing=\begin{smallmatrix}\cdot&&\cdot\\&\cdot&\\\cdot&&\cdot\end{smallmatrix}$.

If $\mathcal{U}=\left\{\begin{smallmatrix}1&&0\\&0&\\0&&0\end{smallmatrix},\ \begin{smallmatrix}0&&1\\&0&\\0&&0\end{smallmatrix},\ \begin{smallmatrix}0&&0\\&1&\\0&&0\end{smallmatrix},\ \begin{smallmatrix}0&&0\\&0&\\1&&0\end{smallmatrix},\ \begin{smallmatrix}0&&0\\&0&\\0&&1\end{smallmatrix}\right\}$ then $\mathcal{U}\subset\mathcal{P}(X)$ but $\mathcal{U}$ fails all the conditions in 1, 2, 3 above:

1) $X=\begin{smallmatrix}1&&1\\&1&\\1&&1\end{smallmatrix}\notin\mathcal{U}$ and $\varnothing=\begin{smallmatrix}0&&0\\&0&\\0&&0\end{smallmatrix}\notin\mathcal{U}$

2) Let $P=\begin{smallmatrix}1&&0\\&0&\\0&&0\end{smallmatrix}\in\mathcal{U}$ and $Q=\begin{smallmatrix}0&&1\\&0&\\0&&0\end{smallmatrix}\in\mathcal{U}$. Then $P\cap Q=\begin{smallmatrix}0&&0\\&0&\\0&&0\end{smallmatrix}\notin\mathcal{U}$ and $P\cup Q=\begin{smallmatrix}1&&1\\&0&\\0&&0\end{smallmatrix}\notin\mathcal{U}$.

3) Let $\mathcal{V}=\left\{\begin{smallmatrix}0&&1\\&0&\\0&&0\end{smallmatrix},\ \begin{smallmatrix}0&&0\\&1&\\0&&0\end{smallmatrix},\ \begin{smallmatrix}0&&0\\&0&\\1&&0\end{smallmatrix}\right\}\subset\mathcal{U}$. Then $\bigcup\mathcal{V}=\begin{smallmatrix}0&&1\\&0&\\0&&0\end{smallmatrix}\cup\begin{smallmatrix}0&&0\\&1&\\0&&0\end{smallmatrix}\cup\begin{smallmatrix}0&&0\\&0&\\1&&0\end{smallmatrix}=\begin{smallmatrix}0&&1\\&1&\\1&&0\end{smallmatrix}\notin\mathcal{U}$.

Now let $K=\begin{smallmatrix}&\bullet&\\\bullet&&\bullet\\&\bullet&\\&\bullet&\end{smallmatrix}$ and $\mathcal{U}=\left\{\begin{smallmatrix}&0&\\0&&0\\&0&\\&0&\end{smallmatrix},\ \begin{smallmatrix}&0&\\0&&0\\&0&\\&1&\end{smallmatrix},\ \begin{smallmatrix}&0&\\0&&0\\&1&\\&1&\end{smallmatrix},\ \begin{smallmatrix}&0&\\0&&1\\&1&\\&1&\end{smallmatrix},\ \begin{smallmatrix}&0&\\1&&0\\&1&\\&1&\end{smallmatrix},\ \begin{smallmatrix}&0&\\1&&1\\&1&\\&1&\end{smallmatrix},\ \begin{smallmatrix}&1&\\1&&1\\&1&\\&1&\end{smallmatrix}\right\}$. In this case $(K,\mathcal{U})$ is a topological space.

Some sets have "default" topologies on them, denoted with '$\mathcal{O}$'. For example, $\mathbb{R}$ is often used to mean the topological space $(\mathbb{R}, \mathcal{O}(\mathbb{R}))$, where:

$$\mathcal{O}(\mathbb{R}) = \{\, U \subset \mathbb{R} \mid U \text{ is a union of open intervals} \,\}.$$

We say that a subset $U \subset \mathbb{R}$ is "open in $\mathbb{R}$" ("in the default sense"; note that now we are saying just "open in $\mathbb{R}$", not "open in $(\mathbb{R}, \mathcal{O}(\mathbb{R}))$") when U is a union of open intervals, i.e., when $U \in \mathcal{O}(\mathbb{R})$; but note that $\mathcal{P}(\mathbb{R})$ and $\{\varnothing, \mathbb{R}\}$ are also topologies on $\mathbb{R}$, and:

$$\begin{array}{lll}
\{2,3,4\} \in \mathcal{P}(\mathbb{R}), & \text{so} & \{2,3,4\} \text{ is open in } (\mathbb{R}, \mathcal{P}(\mathbb{R})), \\
\{2,3,4\} \notin \mathcal{O}(\mathbb{R}), & \text{so} & \{2,3,4\} \text{ is not open in } (\mathbb{R}, \mathcal{O}(\mathbb{R})), \\
\{2,3,4\} \notin \{\varnothing, \mathbb{R}\}, & \text{so} & \{2,3,4\} \text{ is not open in } (\mathbb{R}, \{\varnothing, \mathbb{R}\});
\end{array}$$

when we say just "U is open in X", this means that:
1) $\mathcal{O}(X)$ is clear from the context, and
2) $U \in \mathcal{O}(X)$.

12 The default topology on a ZSet

Let's define a default topology $\mathcal{O}(D)$ for each ZSet D.

For each ZSet D we define $\mathcal{O}(D)$ as:

$$\begin{array}{rcl}
\mathcal{O}(D) & := & \{\, U \subset D \mid \forall ((x,y),(x',y')) \in \mathsf{BPM}(D). \\
 & & \qquad\qquad\qquad (x,y) \in U \to (x',y') \in U \,\}
\end{array}$$

whose visual meaning is this. Turn D into a ZDAG by adding arrows for the black pawns moves (sec.2), and regard each subset $U \subset D$ as a board configuration in which the black pieces may move down to empty positions through the arrows. A subset U is "stable" when no moves are possible because all points of U "ahead" of a black piece are already occupied by black pieces; a subset U is "non-stable" when there is at least one arrow $((x,y),(x',y')) \in \mathsf{BPM}(D)$ in which (x,y) had a black piece and (x',y') is an empty position.

In our two notations for subsets (sec.1) a subset $U \subset D$ is unstable when it has an arrow like '$\bullet \to \cdot$' or '$1 \to 0$'; remember that black pawn moves arrows go down. A subset $U \subset D$ is stable when none of its '$\bullet$'s or '1's can move down to empty positions.

"Open" is the same as "stable". $\mathcal{O}(D)$ is the set of stable subsets of D.

Some examples:

$\begin{smallmatrix} & 0 & \\ 0 & & 1 \\ & 0 & \\ & 0 & \end{smallmatrix}$ is not open because it has a 1 above a 0,

$\mathcal{O}\left(\begin{smallmatrix} & \bullet & \\ \bullet & & \bullet \\ & \bullet & \\ & \bullet & \end{smallmatrix}\right) = \left\{ \begin{smallmatrix} & 0 & \\ 0 & & 0 \\ & 0 & \\ & 0 & \end{smallmatrix}, \begin{smallmatrix} & 0 & \\ 0 & & 0 \\ & 0 & \\ & 1 & \end{smallmatrix}, \begin{smallmatrix} & 0 & \\ 0 & & 0 \\ & 1 & \\ & 1 & \end{smallmatrix}, \begin{smallmatrix} & 0 & \\ 0 & & 1 \\ & 1 & \\ & 1 & \end{smallmatrix}, \begin{smallmatrix} & 0 & \\ 1 & & 0 \\ & 1 & \\ & 1 & \end{smallmatrix}, \begin{smallmatrix} & 0 & \\ 1 & & 1 \\ & 1 & \\ & 1 & \end{smallmatrix}, \begin{smallmatrix} & 1 & \\ 1 & & 1 \\ & 1 & \\ & 1 & \end{smallmatrix} \right\},$

$$\mathcal{O}(\begin{smallmatrix}&\bullet&\\\bullet&&\bullet\\\bullet&&\bullet\end{smallmatrix}) = \left\{ \begin{smallmatrix}&0&\\0&&0\\0&&0\end{smallmatrix}, \begin{smallmatrix}&0&\\0&&0\\0&&1\end{smallmatrix}, \begin{smallmatrix}&0&\\0&&0\\1&&0\end{smallmatrix}, \begin{smallmatrix}&0&\\0&&0\\1&&1\end{smallmatrix}, \begin{smallmatrix}&0&\\0&&1\\0&&1\end{smallmatrix}, \begin{smallmatrix}&0&\\0&&1\\1&&1\end{smallmatrix}, \begin{smallmatrix}&0&\\1&&0\\1&&0\end{smallmatrix}, \begin{smallmatrix}&0&\\1&&0\\1&&1\end{smallmatrix}, \begin{smallmatrix}&0&\\1&&1\\1&&1\end{smallmatrix}, \begin{smallmatrix}&1&\\1&&1\\1&&1\end{smallmatrix} \right\}.$$

The definition of $\mathcal{O}(D)$ above can be generalized to any directed graph. If (A, R) is a directed graph, then $(A, \mathcal{O}_R(A))$ is a topological space if we define:

$$\mathcal{O}_R(A) := \{\, U \subseteq A \mid \forall (a, b) \in R.\ (a \in U \to b \in U) \,\}$$

The two definitions are related as this: $\mathcal{O}(D) = \mathcal{O}_{\mathsf{BPM}(D)}(D)$.

Note that we can see the arrows in $\mathsf{BPM}(D)$ or in R as *obligations* that open sets must obey; each arrow $a \to b$ says that every open set that contains a is forced to contain b too.

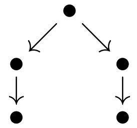
13 Topologies as partial orders

For any topological space $(X, \mathcal{O}(X))$ we can regard $\mathcal{O}(X)$ as a partial order, ordered by inclusion, with $\varnothing$ as its minimal element and X as its maximal element; we denote that partial order by $(\mathcal{O}(X), \subseteq)$.

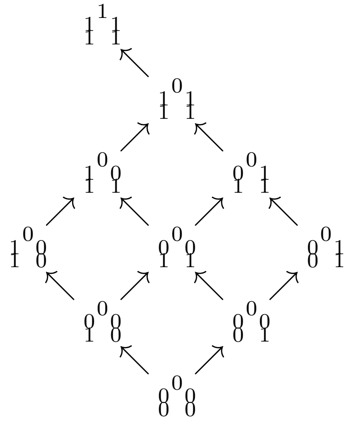
Take any ZSet D. The partial order $(\mathcal{O}(D), \subseteq)$ will *sometimes* be a ZHA when we draw it with $\varnothing$ at the bottom, D at the top, and inclusions pointing up, as can be seen in the three figures below; when $D = \begin{smallmatrix}&\bullet&\\\bullet&&\bullet\\\bullet&&\bullet\end{smallmatrix}$ or $D = \begin{smallmatrix}&\bullet&&\bullet\\\bullet&&\bullet&\\&\bullet&&\bullet\end{smallmatrix}$ the result is a ZHA, but when $D = \begin{smallmatrix}\bullet&&\bullet&&\bullet\\&\bullet&&\bullet&\end{smallmatrix}$ it is not.

Let's write "$V \subset_1 U$" for "$V \subseteq U$ and V and U differ in exactly one point". When D is a ZSet the relation $\subseteq$ on $\mathcal{O}(D)$ is the transitive-reflexive closure of $\subset_1$, and $(\mathcal{O}(D), \subset_1)$ is easier to draw than $(\mathcal{O}(D), \subseteq)$.

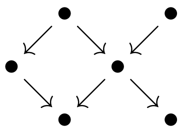
$(H, \mathsf{BPM}(H)) =$



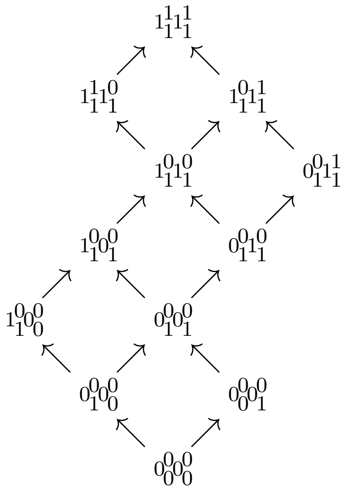
$(\mathcal{O}(H), \subset_1) =$



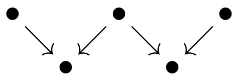
$(G, \mathsf{BPM}(G)) =$



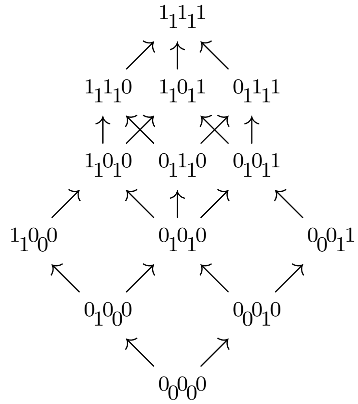
$(\mathcal{O}(G), \subset_1) =$



$(W, \mathsf{BPM}(W)) =$



$(\mathcal{O}(W), \subset_1) =$



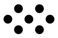
We can formalize a “way to draw $\mathcal{O}(D)$ as a ZHA” (or “...as a ZDAG”) as a bijective function f from a ZHA (or from a ZSet) S to $\mathcal{O}(D)$ that creates a perfect correspondence between the white moves in S and the “$V \subset_1 U$-arrows”; more precisely, an f such that this holds: if $a, b \in S$ then $(a, b) \in \mathsf{WPM}(S)$ iff $f(a) \subset_1 f(b)$.

Note that the *number of elements* in an open set corresponds to the *height* where it is drawn; if $f : S \to \mathcal{O}(D)$ is a way to draw $\mathcal{O}(D)$ as a ZHA or a ZDAG then f takes points of the form $(\_\_\_, y)$ to open sets with y elements, and if $f : S \to \mathcal{O}(D)$ is a way to draw $\mathcal{O}(D)$ as a ZHA (not a ZDAG!) then we also have that $f((0, 0)) = \varnothing \in \mathcal{O}(D)$.

The diagram for $(\mathcal{O}(H), \subset_1)$ above is a way to draw $\mathcal{O}(H)$ as a ZHA.

The diagram for $(\mathcal{O}(G), \subset_1)$ above is a way to draw $\mathcal{O}(G)$ as a ZHA.

The diagram for $(\mathcal{O}(W), \subset_1)$ above is *not* a way to draw $\mathcal{O}(W)$ as a ZSet. Look at ${}^{0}{}_{1}{}^{1}{}_{1}{}^{0}$ and ${}^{1}{}_{1}{}^{0}{}_{1}{}^{1}$ in the middle of the cube formed by all open sets of the form ${}^{a}{}_{1}{}^{b}{}_{1}{}^{c}$. We don’t have ${}^{0}{}_{1}{}^{1}{}_{1}{}^{0} \subset_1 {}^{1}{}_{1}{}^{0}{}_{1}{}^{1}$, but we do have a white pawn move (not draw in the diagram!) from $f^{-1}({}^{0}{}_{1}{}^{1}{}_{1}{}^{0})$ to $f^{-1}({}^{1}{}_{1}{}^{0}{}_{1}{}^{1})$. We say that a ZSet is *thin* when it doesn’t have three independent points.

Every time that a ZSet D has three independent points, as in W, we will have a situation like in $(\mathcal{O}(W), \subset_1)$; for example, if $B =$  then the open sets of B of the form $a^{0}_{1}b^{0}_{1}c$ form a cube.

14 2-Column Graphs

Note: in this section we will manipulate objects with names like $1\_, 2\_, 3\_, \ldots, \_1, \_2, \_3, \ldots$; here are two good ways to formalize them:

$$\begin{array}{ll}
\vdots & \vdots \\
4\_ = (0,4) & \_4 = (1,4) \\
3\_ = (0,3) & \_3 = (1,3) \\
2\_ = (0,2) & \_2 = (1,2) \\
1\_ = (0,1) & \_1 = (1,1)
\end{array}
\quad \text{or} \quad
\begin{array}{ll}
\vdots & \vdots \\
4\_ = \texttt{"4\_"} & \_4 = \texttt{"\_4"} \\
3\_ = \texttt{"3\_"} & \_3 = \texttt{"\_3"} \\
2\_ = \texttt{"2\_"} & \_2 = \texttt{"\_2"} \\
1\_ = \texttt{"1\_"} & \_1 = \texttt{"\_1"}
\end{array},$$

where `"1_"`, `"_2"`, `""`, `"Hello!"`, etc are strings.

We define:

$$\begin{array}{rcl}
\mathsf{LC}(l) & := & \{1\_, 2\_, \ldots, l\_\} \\
\mathsf{RC}(r) & := & \{\_1, \_2, \ldots, \_r\},
\end{array}$$

which generate a “left column” of height l and a “right column” of height r.

A *description for a 2-column graph* (a “D2CG”) is a 4-tuple (l, r, R, L), where $l, r \in \mathbb{N}$, $R \subset \mathsf{LC}(l) \times \mathsf{RC}(r)$, $L \subset \mathsf{RC}(r) \times \mathsf{LC}(l)$; l is the height of the left column, r is the height of the right column, and R and L are set of intercolumn arrows (going right and left respectively).

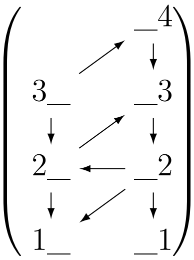
The operation 2CG (in a sans-serif font) generates a directed graph from a D2CG:

$$\mathsf{2CG}(l,r,R,L) \quad := \quad \left(\mathsf{LC}(l) \cup \mathsf{RC}(r), \left\{ \begin{array}{c} \{l\_{\to}(l{-}1)\_, \ldots, 2\_{\to}1\_\}\cup \\ \{\_r{\to}\_(r{-}1), \ldots, \_2{\to}\_1\}\cup \\ R\cup L \end{array} \right\} \right)$$

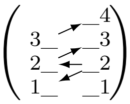
For example,

$$\mathsf{2CG}(3,4,\left\{ \begin{array}{c} 3\_{\to}\_4, \\ 2\_{\to}\_3 \end{array} \right\}, \left\{ \begin{array}{c} 2\_{\leftarrow}\_2, \\ 1\_{\leftarrow}\_2 \end{array} \right\}) \quad := \quad \left(\left\{ \begin{array}{c} 3\_, 2\_, 1\_, \\ \_4, \_3, \_2, \_1 \end{array} \right\}, \left\{ \begin{array}{c} 3\_{\to}2\_, 2\_{\to}1\_, \\ \_4{\to}\_3, \_3{\to}\_2, \_2{\to}\_1, \\ 3\_{\to}\_4, 2\_{\to}\_3, \\ 2\_{\leftarrow}\_2, 1\_{\leftarrow}\_2 \end{array} \right\} \right)$$

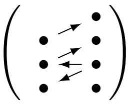
which is:



we will usually draw that more compactly, by omitting the intracolumn (i.e., vertical) arrows:



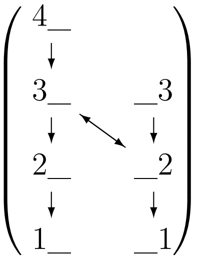
or



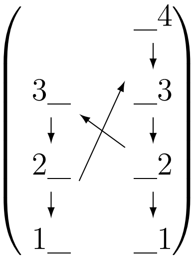
.

A *2-column graph* (a "2CG") is a directed graph that is of the form $\mathsf{2CG}(l,r,R,L)$. We will often say $(P,A) = \mathsf{2CG}(l,r,R,L)$, where the P stand for "points" and A for "arrows".

A *2-column acyclical graph* (a "2CAG") is a 2CG that doesn't have cycles. If L has an arrow that is the opposite of an arrow in R, this generates a cycle of length 2; if R has an arrow $l\_{\to}\_r'$ and L has an arrow $l'\_{\leftarrow}\_r$, where $l \leq l'$ and $r \leq r'$, this generates a cycle that can have a more complex shape — a triangle or a bowtie. For example,



and



.

15 Topologies on 2CGs

In this section we will see that ZHAs are topologies on 2CAGs.

Let $(P, A) = \mathsf{2CG}(l, r, R, L)$ be a 2-column graph.
What happens if we look at the open sets of (P, A), i.e., at $\mathcal{O}_A(P)$? Two things:
1) every open set $U \in \mathcal{O}_A(P)$ is of the form $\mathsf{LC}(a) \cup \mathsf{RC}(b)$,
2) arrows in R and L forbids some '$\mathsf{LC}(a) \cup \mathsf{RC}(b)$'s from being open sets.
In order to understand that we need to introduce some notations for "piles".

The function

$$\mathsf{pile}(\langle a, b\rangle) := \mathsf{LC}(a) \cup \mathsf{RC}(b)$$

converts an element $\langle a, b\rangle \in \mathbb{LR}$ into a pile of elements in the left column of height a and a pile of elements in the right column of height b. We will write subsets of the points of a 2CG using a positional notation with arrows. So, for example, if $(P, A) = \mathsf{2CG}(3, 4, \{2\_\!\to\!\_3\}, \{2\_\!\leftarrow\!\_2\})$ then

$$(P, A) = \begin{pmatrix} & & \_4 \\ 3\_ & \nearrow & \_3 \\ 2\_ & \leftarrow & \_2 \\ 1\_ & & \_1 \end{pmatrix} \qquad \text{and} \qquad \mathsf{pile}(21) = \begin{pmatrix} & & 0 \\ 0 & \nearrow & 0 \\ 1 & \leftarrow & 0 \\ 1 & & 1 \end{pmatrix} \ \text{(as a subset of } P\text{).}$$

Note that $\mathsf{pile}(21)$ is not open in $(P, \mathcal{O}_A(P))$, as it has an arrow '$1 \to 0$'. In fact, the presence of the arrow $\{2\_\!\to\!\_3\}$ in A means that all piles of the form

$$\begin{pmatrix} & & 0 \\ ? & \nearrow & 0 \\ 1 & \leftarrow & ? \\ 1 & & ? \end{pmatrix}$$

are not open, the presence of the arrow $\{2\_\!\leftarrow\!\_2\}$ means that the piles of the form

$$\begin{pmatrix} & & ? \\ 0 & \nearrow & ? \\ 0 & \leftarrow & 1 \\ ? & & 1 \end{pmatrix}$$

are not open sets.

The effect of these prohibitions can be expressed nicely with implications. If

$$(P, A) = \mathsf{2CG}\left(l, r, \left\{ \begin{smallmatrix} c\_\to\_d, \\ e\_\to\_f \end{smallmatrix} \right\}, \left\{ \begin{smallmatrix} g\_\leftarrow\_h, \\ i\_\leftarrow\_j \end{smallmatrix} \right\}\right)$$

then

$$\mathcal{O}_A(P) = \left\{ \mathsf{pile}(ab) \mid a \in \{0, \dots, l\}, b \in \{0, \dots, r\}, \begin{pmatrix} a{\geq}c \to b{\geq}d \,\land \\ a{\geq}e \to b{\geq}f \,\land \\ a{\geq}g \leftarrow b{\geq}h \,\land \\ a{\geq}i \leftarrow b{\geq}j \end{pmatrix} \right\}$$

Let’s use a shorter notation for comparing 2CGs and their topologies:

$$\mathcal{O}\left(\begin{tikzcd}[row sep=small] & \_5 \arrow[d] \\ 4\_ \arrow[d] & \_4 \arrow[d] \arrow[dl] \\ 3\_ \arrow[d] \arrow[dr] & \_3 \arrow[d] \arrow[dl] \\ 2\_ \arrow[d] & \_2 \arrow[d] \\ 1\_ & \_1 \end{tikzcd}\right) = \begin{array}{cccccc} &&&45&& \\ &&44&&35& \\ &43&&34&&25 \\ 42&&33&&24& \\ &32&&23&&14 \\ &&22&&13& \\ &21&&12&&03 \\ 20&&11&&02& \\ &10&&01&& \\ &&00&&& \end{array}$$

the arrows in R and L and the values of l and r are easy to read from the 2CG at the left, and we omit the ‘pile’s at the right.

In a situation like the above we say that the 2CG in the ‘$\mathcal{O}(\ldots)$’ *generates* the ZHA at the right. There is an easy way to draw the ZHA generated by a 2CG, and a simple way to find the 2CG that generates a given ZHA. To describe them we need two new concepts.

If (A,R) is a directed graph and $S \subset A$ then ${\downarrow}S$ is the smallest open set in $\mathcal{O}_R(A)$ that contains S. If (A,R) is a ZDAG with black pawns moves as its arrows, think that the ‘1’s in S are painted with a black paint that is very wet, and that that paint flows into the ‘0’s below; the result of ${\downarrow}S$ is what we get when all the ‘0’s below ‘1’s get painted black. For example: ${\downarrow}\,0{}^{0}_{0}0{}^{1}_{0} = 0{}^{0}_{1}1{}^{1}_{1}$. When (P,A) is a 2CG and $S \subseteq P$, we have to think that the paint flows along the arrows, even if some of the intercolumn arrows point upward. For example:

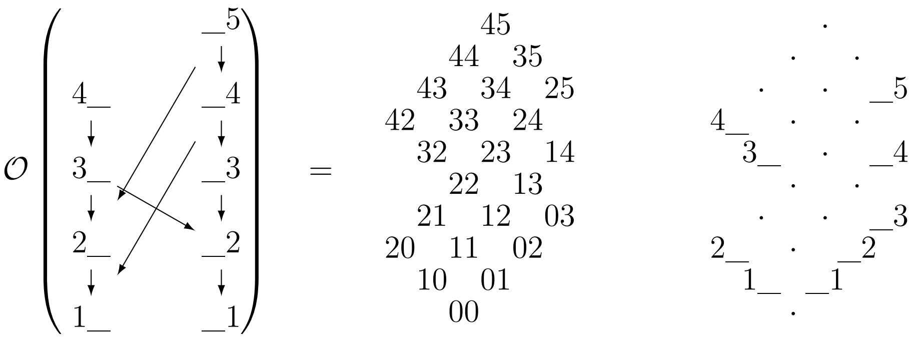
$${\downarrow}\left(\begin{tikzcd}[row sep=tiny, column sep=small] & 0 \\ 0 & 0 \\ 0 \arrow[ur] & 1 \arrow[l] \\ 1 & 0 \end{tikzcd}\right) = \left(\begin{tikzcd}[row sep=tiny, column sep=small] & 0 \\ 0 & 1 \\ 1 \arrow[ur] & 1 \arrow[l] \\ 1 & 1 \end{tikzcd}\right)$$

and if S consists of a single point, $S = \{s\}$, then we may write ${\downarrow}s$ instead of ${\downarrow}\{s\} = {\downarrow}S$. In the 2CG above, we have (omitting the ‘pile’s):

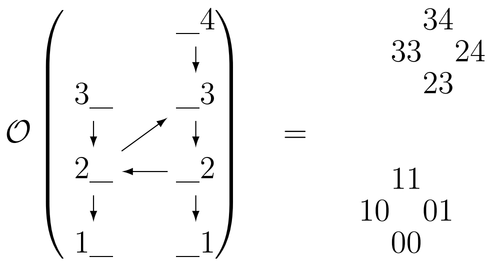
$${\downarrow}\_2 \;=\; {\downarrow}\{\_2\} = {\downarrow}\left(\begin{tikzcd}[row sep=tiny, column sep=small] & 0 \\ 0 & 0 \\ 0 \arrow[ur] & 1 \arrow[l] \\ 0 & 0 \end{tikzcd}\right) = \left(\begin{tikzcd}[row sep=tiny, column sep=small] & 0 \\ 0 & 1 \\ 1 \arrow[ur] & 1 \arrow[l] \\ 1 & 1 \end{tikzcd}\right) = 23, \quad \text{and} \quad \begin{array}{ll} & {\downarrow}\_4{=}24, \\ {\downarrow}3\_{=}33, & {\downarrow}\_3{=}23, \\ {\downarrow}2\_{=}23, & {\downarrow}\_2{=}23, \\ {\downarrow}1\_{=}10, & {\downarrow}\_1{=}01, \end{array}$$

The second concept is this: the “generators” of a ZDAG D with white pawns moves as its arrows — or of a ZHA D — are the points of D that have exactly one white pawn move pointing *to* them (not *going out of* them).

If (P,A) is a 2CAG, then $\mathcal{O}_A(P)$ is a ZHA, and ‘${\downarrow}$’ is a bijection from P to the generators of $\mathcal{O}_A(P)$; for example:



but if (P, A) is a 2CG with cycles, then $\mathcal{O}_A(P)$ is not a ZHA because each cycle generates a "gap" that disconnects the points of $\mathcal{O}_A(P)$. We just saw an example of a 2CG with a cycle in which $\downarrow 2\_ = 23 = \downarrow \_3 = \downarrow \_2$; look at its topology:



16 Topologies as Heyting Algebras

The *open-set semantics* for Intuitionistic Propositional Logic is based on this idea: choose any topological space $(X, \mathcal{O}(X))$; the opens sets of $\mathcal{O}(X)$ will play the role of truth-values, and we define the components of a Heyting Algebra (sec.8) as this:

$$\begin{array}{rcl@{\qquad}cl}
\Omega & := & \mathcal{O}(X) & & \\
P \leq Q & := & P \subseteq Q & & \\
\top & := & \{\, x \in X \mid \top \,\} & = & X \\
\bot & := & \{\, x \in X \mid \bot \,\} & = & \emptyset \\
P \land Q & := & \{\, x \in X \mid x \in P \land x \in Q \,\} & = & P \cap Q \\
P \lor Q & := & \{\, x \in X \mid x \in P \lor x \in Q \,\} & = & P \cup Q \\
P \overset{\mathrm{M}}{\to} Q & := & \{\, x \in X \mid x \in P \to x \in Q \,\} & & \\
 & = & \{\, x \in X \mid x \notin P \lor x \in Q \,\} & = & (X \backslash P) \cup Q
\end{array}$$

However, this '$\overset{\mathrm{M}}{\to}$' *may* return a non-open result even when given open inputs,

$$\begin{smallmatrix} & 0 & \\ 1 & & 0 \\ 1 & & 0 \end{smallmatrix} \overset{\mathrm{M}}{\to} \begin{smallmatrix} & 0 & \\ 0 & & 0 \\ 1 & & 1 \end{smallmatrix} = \begin{smallmatrix} & 1 & \\ 0 & & 1 \\ 1 & & 1 \end{smallmatrix}$$

so our definition is broken; we can fix it by taking the interior:

$$P \to Q \quad := \quad \mathsf{int}(P \overset{\mathrm{M}}{\to} Q) \quad = \quad \mathsf{int}((X \backslash P) \cup Q)$$

Theorem 16.1 *For any topological space* $(X, \mathcal{O}(X))$ *the structure* $(\Omega, \le, \top, \bot, \wedge, \vee, \to)$ *defined as above is a Heyting Algebra. In particular, this holds for any* $P, Q, R \in \Omega$*:* $P \le (Q \to R)$ *iff* $(P \wedge Q) \le R$*.*

Proof. Standard; see for example [Awo06] (section 6.3). ■

Note that Theorem 16.1 gives us another way to calculate the connectives in 2CGs. In sec.7 we saw how to calculate $\neg\neg P \to P$ in a certain ZHA when $P = 10$; compare it with the "topological" way, in which the truth-values are subsets of $\begin{smallmatrix} & \bullet & \\ \bullet & & \bullet \\ \bullet & & \bullet \end{smallmatrix}$:

$$\underbrace{(\underbrace{\neg \underbrace{\neg \underbrace{P}_{10}}_{02}}_{20}) \to \underbrace{P}_{10}}_{12} \qquad \underbrace{(\underbrace{\neg \underbrace{\neg \underbrace{P}_{\begin{smallmatrix} & 0 & \\ 0 & & 0 \\ 1 & & 0 \end{smallmatrix}}}_{\begin{smallmatrix} & 0 & \\ 0 & & 1 \\ 0 & & 1 \end{smallmatrix}}}_{\begin{smallmatrix} & 0 & \\ 1 & & 0 \\ 1 & & 0 \end{smallmatrix}}) \to \underbrace{P}_{\begin{smallmatrix} & 0 & \\ 0 & & 0 \\ 1 & & 0 \end{smallmatrix}}}_{\begin{smallmatrix} & 0 & \\ 0 & & 1 \\ 1 & & 1 \end{smallmatrix}}$$

17 Converting between ZHAs and 2CAGs

Let's now see how to start from a 2CAG and produce its topology (a ZHA) quickly, and how to find quickly the 2CAG that generates a given ZHA.

From 2CAGs to ZHAs. Let $(P, A) = \mathsf{2CG}(l, r, R, L)$ be a 2CAG, and call the ZHA generated by it H. Then the top point of H is lr, and its bottom point is 00. Let $C := \{00, \downarrow\!1\_, \downarrow\!2\_, \ldots, \downarrow\!l\_, lr\}$, i.e., the left generators (see the end of sec.15) plus $\bot$ and $\top$; then C has some of the points of the left wall (sec.4) of H, but usually not all. To "complete" C, apply this operation repeatedly: if $ab \in C$ and $ab \neq lr$, then test if either $(a+1)b$ or $a(b+1)$ are in C; if none of them are, add $a(b+1)$, which is northeast of ab. When there is nothing else to add, then C is the whole of the left wall of H. For the right wall, start with $D := \{00, \downarrow\!\_1, \downarrow\!\_2, \ldots, \downarrow\!\_r, lr\}$, and for each $ab \in C$ with $ab \neq lr$, test if either $(a+1)b$ or $a(b+1)$ are in D; if none of them are, add $(a+1)b$, which is northwest of ab. When there is nothing else to add, then D is the whole of the right wall of H.

In the acyclic example of the last section this yields:

$$\begin{array}{rcl} C & = & \{00, {\downarrow}1\_, {\downarrow}2\_, {\downarrow}3\_, {\downarrow}4\_, lr\} \\ & = & \{00, 10, 20, 32, 42, 45\} \\ & \rightsquigarrow & \{00, 10, 20, 21, 22, 32, 42, 43, 44, 45\}, \\ D & = & \{00, {\downarrow}\_1, {\downarrow}\_2, {\downarrow}\_3, {\downarrow}\_4, {\downarrow}\_5, lr\} \\ & = & \{00, 01, 02, 03, 14, 25, 45\} \\ & \rightsquigarrow & \{00, 01, 02, 03, 13, 14, 24, 25, 35, 45\}. \end{array}$$

and the ZHA is everything between the "left wall" C and the "right wall" D.

From ZHAs to 2CAGs. Let H be a ZHA and let lr be its top point. Form the sequence of its left wall generators (the generators of H in which the arrow pointing to them points northwest) and the sequence of its right wall generators (the generators of H in which the arrow pointing to them points northeast). Look at where there are "gaps" in these sequences; each gap in the left wall generators becomes an intercolumn arrow going right, and each gap in the right wall generators becomes an intercolun arrow going left. In the acyclic example of the last section, this yields:

$$\begin{array}{llll} & & \_5 = 25 & \\ & & & \text{(gap becomes } 2\_{\leftarrow}\_5) \\ 4\_ = 42 & & \_4 = 14 & \\ & \text{(no gap)} & & \text{(gap becomes } 1\_{\leftarrow}\_4) \\ 3\_ = 32 & & \_3 = 03 & \\ & \text{(gap becomes } 3\_{\to}\_2) & & \text{(no gap)} \\ 2\_ = 20 & & \_2 = 02 & \\ & \text{(no gap)} & & \text{(no gap)} \\ 1\_ = 10 & & \_1 = 01 & \end{array}$$

We know l and r from the top point of the ZHA, and from the gaps we get R and L; the 2CAG that generates this ZHA is:

$$(4, 5, \{3\_{\to}\_2\}, \left\{\begin{matrix} 2\_{\leftarrow}\_5, \\ 1\_{\leftarrow}\_4 \end{matrix}\right\}).$$

Theorem 17.1 *The two operations above are inverse to one another in the following sense. If we start with a ZHA H, produce its 2CAG, and produce a ZHA H' from that, we get the same ZHA: $H' = H$. In the other direction, if we start with a 2CAG $(P, A) = \mathsf{2CG}(l, r, R, L)$, produce its ZHA, H, and then obtain a 2CAG $(P', A') = \mathsf{2CG}(l', r', R', L')$ from H, we get back the original 2CAG* if and only if it didn't have *any superfluous arrows; if the original 2CAG had superflous arrows then then new 2CAG will have $l' = l$, $r' = r$, and R' and L' will be R and L minus these "superfluous arrows", that are the ones that can be deleted without changing which 2-piles are forbidden. For example:*

$$\begin{tikzcd}
4\_ \arrow[r] \arrow[d] \arrow[dr] & \_4 \arrow[d] \\
3\_ \arrow[d] \arrow[dr] & \_3 \arrow[d] \\
2\_ \arrow[r] \arrow[d] & \_2 \arrow[d] \\
1\_ & \_1
\end{tikzcd}$$

$$\rightsquigarrow \quad \begin{array}{cccccc}
 & 44 & & & & \\
 & & 34 & & & \\
 & 33 & & 24 & & \\
32 & & 23 & & 14 & \\
 & 22 & & 13 & & 04 \\
 & & 12 & & 03 & \\
 & 11 & & 02 & & \\
10 & & 01 & & & \\
 & 00 & & & &
\end{array} \quad \rightsquigarrow$$

$$\begin{tikzcd}
4\_ \arrow[r] \arrow[d] & \_4 \arrow[d] \\
3\_ \arrow[d] & \_3 \arrow[d] \\
2\_ \arrow[r] \arrow[d] & \_2 \arrow[d] \\
1\_ & \_1
\end{tikzcd}$$

In this case we have $R = \left\{ \begin{smallmatrix} 4\_ \to \_4, \\ 4\_ \to \_3, \\ 3\_ \to \_2, \\ 2\_ \to \_2 \end{smallmatrix} \right\}$ *and* $R' = \left\{ \begin{smallmatrix} 4\_ \to \_4, \\ 2\_ \to \_2 \end{smallmatrix} \right\}$.

18 ZHA Logic is between IPL and CPL

In standard terminology, this is: ZHA Logic is a superintuitionistic logic ([CZ97], p.109) of "bounded width 2", i.e., where the axiom $\mathbf{BW}_2$ of [CZ97], p.112, holds. But let's see this in elementary terms.

Let S be this sentence:

$$\begin{array}{rcl}
S_P & := & P \to (Q \lor R) \\
S_Q & := & Q \to (R \lor P) \\
S_R & := & R \to (P \lor Q) \\
S & := & S_P \lor S_Q \lor S_R
\end{array}$$

S can't be an intuitionistic theorem because in this Heyting Algebra, with these values for P, Q, R,

$$(W, A) = \begin{tikzcd}
 & 1 \arrow[dl] \arrow[d] \arrow[dr] & \\
2 & 3 & 4
\end{tikzcd}
\qquad
(\mathcal{O}_A(W), \subset_1) = \begin{tikzcd}
 & 1\overset{1}{1}1 & \\
 & 1\overset{0}{1}1 \arrow[u] & \\
1\overset{0}{1}0 \arrow[ur] & 1\overset{0}{0}1 \arrow[u] & 0\overset{0}{1}1 \arrow[ul] \\
1\overset{0}{0}0 \arrow[u] \arrow[ur] & 0\overset{0}{1}0 \arrow[ul] \arrow[ur] & 0\overset{0}{0}1 \arrow[ul] \arrow[u] \\
 & 0\overset{0}{0}0 \arrow[ul] \arrow[u] \arrow[ur] &
\end{tikzcd}
\qquad
\begin{array}{rcl}
P & = & 1\overset{0}{0}0 \\
Q & = & 0\overset{0}{1}0 \\
R & = & 0\overset{0}{0}1
\end{array}$$

we have $S = 1\overset{0}{1}1 \neq \top = 1\overset{1}{1}1$.

One way to define a *valuation* for a sentence S with variables $\mathsf{Vars}(S)$ — in our example we have $\mathsf{Vars}(S) = \{P, Q, R\}$) — is as a pair made of a Heyting Algebra H and a function $v : \mathsf{Vars}(S) \to H$. A looser definition is that a valuation for S is a pair made of 1) something that generates a Heyting Algebra in a known, canonical way, and 2) a function from $\mathsf{Vars}(S)$ to the elements of that HA. So:

A *classical valuation* for S is a valuation of the form $(\{0, 1\}, v)$.

A *ZHA-valuation* for S is a valuation of the form (H, v), where H is a ZHA.

A *finite DAG-valuation* for S is a valuation of the form $((W, A), v)$, where W is a finite set and $A \subseteq W \times W$ is a set of arrows on W; the Heyting Algebra on $(W, \mathcal{O}_A(W))$ is built as in sec.16.

A *2CG-valuation* for S is a finite DAG-valuation for S of the form $((P, A), v)$, where (P, A) is a 2-column graph; each 2CG-valuation is naturally equivalent to a ZHA-valuation, and vice-versa.

A *classical countermodel* for S is classical valuation for S in which the value of S is not $\top$; a *ZHA-countermodel* for S is a ZHA-valuation for S in which the value of S is not $\top$; an *intuitionistic countermodel* for S is a finite DAG-valuation for S in which the value of S is not $\top$.

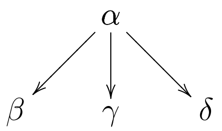
A sentence S is a *classical tautology* (notation: $S \in \mathsf{Taut}(\mathsf{CPL})$) if S has no classical countermodels; a sentence S is a *ZHA-tautology* (notation: $S \in \mathsf{Taut}(\mathsf{ZHAL})$); and a sentence S is an *intuitionistic tautology* (notation: $S \in \mathsf{Taut}(\mathsf{IPL})$) of S has no finite-DAG countermodels.

It is a standard result that the intuitionistic *theorems* are exactly the finite-DAG *tautologies*; this can be seen using Gödel translation (see [Göd86] and [Tro86]) to translate S to S4, and then using modal tableaux for S4 ([Fit72]) to look for a countermodel; in standard terminology, W is a set of "worlds", A is an "accessibility relation" or a notion of which worlds are "ahead" of which other ones, and (W, A^*) is a Kripke frame for S4.

The sentence $S = S_P \vee S_Q \vee S_R$ of the beginning of the section is a good example for introducting tableau methods for modal logics to "children", as the tableau that it generates doesn't have branches. We can present the method directly and in elementary terms, as we will do now.

Fix a set W and a relation $A \subseteq W \times W$. We will say that β is "ahead" of α when $(\alpha, \beta) \in A^*$, i.e., when there is a path $\alpha \to \ldots \to \beta$ using only arrows in A. Let P and Q be open sets in $\mathcal{O}_A(W)$. The only way to have $P \vee Q$ false in a world α (notation: $(P \vee Q)_\alpha = 0$) is to have $P_\alpha = 0$ and $Q_\alpha = 0$. The only way to have $P \to Q$ false in a world α, i.e., $(P \to Q)_\alpha = 0$ is to have $P_\beta = 1$ and $Q_\beta = 0$ in *some* world β, with β ahead of α.

Let $((W, A), v)$ be a finite DAG-countermodel for $S = S_P \vee S_Q \vee S_R$. Then $v(P)$, $v(Q), v(R) \in \mathcal{O}_A(W)$; we will omit the '$v$'s. If $((W, A), v)$ is a countermodel this means that $S \neq \top$, and there is some world α in W in which $S_\alpha = 0$. Fix this α. $S_\alpha = 0$ means $(S_P \vee S_Q \vee S_R)_\alpha = 0$, which means that $(S_P)_\alpha = 0$, $(S_Q)_\alpha = 0$, and $(S_R)_\alpha = 0$. $(S_P)_\alpha = 0$ means $(P \to (Q \vee R))_\alpha = 0$, which means that there is a world β ahead of α in which $P_\beta = 1$ and $(Q \vee R)_\beta = 0$, and $(Q \vee R)_\beta = 0$ means $Q_\beta = 0$ and $R_\beta = 0$; similarly, $(S_Q)_\alpha = 0$ means that there is a world γ ahead of α in which $Q_\gamma = 1$, $R_\gamma = 0$, $P_\gamma = 0$, and $(S_R)_\alpha = 0$ means that there is a world δ ahead of α in which $R_\delta = 1$, $P_\delta = 0$, $Q_\delta = 0$. In diagrams:



$$\begin{tikzcd}
& \begin{array}{c} S_\alpha = 0 \\ (S_P)_\alpha = (P \to (Q \vee R))_\alpha = 0 \\ (S_Q)_\alpha = (Q \to (R \vee P))_\alpha = 0 \\ (S_R)_\alpha = (R \to (P \vee Q))_\alpha = 0 \end{array} \arrow[dl] \arrow[d] \arrow[dr] & \\
\begin{array}{c} P_\beta = 1 \\ (Q \vee R)_\beta = 0 \\ Q_\beta = 0 \\ R_\beta = 0 \end{array} & \begin{array}{c} Q_\gamma = 1 \\ (R \vee P)_\gamma = 0 \\ R_\gamma = 0 \\ P_\gamma = 0 \end{array} & \begin{array}{c} R_\delta = 1 \\ (P \vee Q)_\delta = 0 \\ P_\delta = 0 \\ Q_\delta = 0 \end{array}
\end{tikzcd}$$

Note that β and γ are "independent" in the sense that in A^* we can't have an arrow $\beta \to \gamma$ and neither an arrow $\gamma \to \beta$; we can't have $\beta \to \gamma$ because $P_\beta = 1$ but $P_\gamma = 0$, and we can't have $\gamma \to \beta$ because $Q_\gamma = 1$ but $Q_\beta = 0$. We can use a similar argument to show that γ and δ are independent, and to show also that δ and β are independent.

We can't have three independent points in a 2-column graph, so we have finite DAG-countermodels for S but no 2CG-countermodels for S, and so no ZHA-countermodels for S. This means that S is not an intuitionistic tautology, but it is a ZHA-tautology. It is easy to see that $\mathsf{Taut(IPL)} \subset \mathsf{Taut(ZHAL)} \subset \mathsf{Taut(CPL)}$, and we saw that $S \notin \mathsf{Taut(IPL)}$, $S \in \mathsf{Taut(ZHAL)}$, $(\neg\neg P \to P) \notin \mathsf{Taut(ZHAL)}$, $(\neg\neg P \to P) \in \mathsf{Taut(IPL)}$, which means that:

$$\mathsf{Taut(IPL)} \subsetneq \mathsf{Taut(ZHAL)} \subsetneq \mathsf{Taut(CPL)}$$

and so "ZHA Logic", *which we have not defined via a deduction system*, only by the notions of "ZHA countermodels" and "ZHA tautologies", is strictly between Intuitionistic Logic and Classical Logic, and is different from both.

References

[Awo06] S. Awodey. *Category Theory*. Oxford University Press, 2006.

[Cor04] D. Corfield. *Towards a Philosophy of Real Mathematics*. Cambridge, 2004.

[CZ97] A. Chagrov and M. Zakharyaschev. *Modal Logic*. Oxford, 1997.

[Dal08] D. van Dalen. *Logic and Structure (4th ed.)* Springer, 2008.

[Fit72] M. Fitting. "Tableau Methods of Proof For Modal Logics". In: *Notre Dame Journal of Formal Logic* XIII.2 (Apr. 1972), pp. 237–247.

[Göd86] K. Gödel. "An interpretation of the intuitionistic propositional calculus (1933f)". In: *Kurt Gödel, Collected Works vol.1 (Publications 1929–1936).* Ed. by S. Feferman. Oxford, 1986, p. 301.

[Gol84] R. Goldblatt. *Topoi: The Categorial Analysis of Logic.* North Holland, 1984.

[Joh77] P. T. Johnstone. *Topos Theory.* Academic Press, 1977.

[Krö07] R. Krömer. *Tool and Object: A History and Philosophy of Category Theory.* Birkhäuser, 2007.

[NP01] S. Negri and J. von Plato. *Structural Proof Theory.* Cambridge, 2001.

[Och13] E. Ochs. "Internal Diagrams and Archetypal Reasoning in Category Theory". In: *Logica Universalis* 7.3 (Sept. 2013), pp. 291–321.

[Och18] E. Ochs. "Planar Heyting Algebras for Children 2: Closure Operators". `http://angg.twu.net/math-b.html\#zhas-for-children-2`. 2018.

[Och19a] E. Ochs. "A diagram for the Yoneda Lemma (In which each node and arrow can be interpreted precisely as a "term", and most of the interpretations are "obvious"; plus dictionaries!!!)" `http://angg.twu.net/LATEX/2019notes-yoneda.pdf`. 2019.

[Och19b] E. Ochs. "On some missing diagrams in the Elephant". `http://angg.twu.net/LATEX/2019oxford-abs.pdf`. 2019.

[Tro86] A.S. Troelstra. "Introductory note to 1933f". In: *Kurt Gödel, Collected Works vol.1 (Publications 1929–1936).* Ed. by S. Feferman. Oxford, 1986, pp. 296–299.

Eduardo Ochs
Departmento de Ciências da Natureza
Universidade Federal Fluminense (UFF)
Rua Recife, Lotes 1-7, Jardim Bela Vista, Rio das Ostras, RJ, CEP 28895-532, Brazil
E-mail: `eduardoochs@gmail.com`