

Master WIP

Kjetil Midtgarden Golid

March 9, 2016

Contents

1	Introduction	2
1.1	Paradoxes	2
1.2	Discourses	3
1.3	Graphs, Kernels and Solutions	4
1.4	Discourse Theories and Digraphs	5
1.5	Results in Kernel Theory	7
1.6	Recognizing dags without kernels	9
1.7	Resolving GNF-theories	9
1.7.1	The inference system	11
1.7.2	Inconsistency of the Yablo-graph	12
1.8	This thesis	12
1.8.1	Motivation	12
1.8.2	Thesis Overview	13
2	NAND-clauses in graphs	14
2.1	Unary and binary NAND-clauses	14
3	Expressional power of binary NAND-clauses	16
3.1	Proving inconsistency of unrestricted theories	16
3.1.1	Proving the pigeonhole principle - 4 pigeons, 3 holes	17
4	Proofs	18
4.1	Translating CNF to GNF	18
4.2	Inconsistency of Stretched Yablo	20
4.2.1	Proof, D1 and D1b	21
4.2.2	Proof, D2	22
4.2.3	Proof, D3	23
4.2.4	Proof, \emptyset	23

Chapter 1

Introduction

1.1 Paradoxes

A theory in propositional logic is semantically *inconsistent* if it has no model, i.e., there exists no variable assignment making all formulae in the theory true. Consider the following example:

$$x \wedge \neg x \tag{1.1}$$

While a sentence like $\neg a \rightarrow b$ can be satisfied by, for instance, letting both a and b be true, no such assignment can be made for the statement above. The statement is therefore inconsistent.

A paradox is usually informally defined as something along the lines of “*a statement that can be neither true nor false*”. We can immediately note one thing from this intuitive definition: Since no paradoxes can be true, all paradoxes are, by definition, inconsistent. It does, however, seem like a stretch to say that all inconsistent theories are paradoxes. Just consider the inconsistent statement, $x \wedge \neg x$ again: this statement simply seems false, and not paradoxical.

A different view is that a paradox is a *dialetheia*, a sentence that is *both* true and false[1]. We will however not spend much time exploring these philosophical differences, as this is not a philosophical paper and it won’t change much for our definitions.

The liar sentence is probably the most famous example of a paradox:

“This sentence is false.”

If the statement is true, then the statement is false, but if the statement is false, then the statement is true. It can thus neither be true nor false. Note how the liar sentence is a statement about other statements (in this case itself). In order to study these kinds of meta-statements, we need a way to reference other statements within a statement.

In propositional logic, we can do this by giving statements "names" in the form of adding fresh variables with equivalences to their corresponding statements¹. Consider the left statements below, together with their corresponding named statements on the right.

$$a \qquad x_1 \leftrightarrow a \qquad (1.2)$$

$$a \wedge \neg a \qquad x_2 \leftrightarrow a \wedge \neg a \qquad (1.3)$$

$$a \vee \neg a \qquad x_3 \leftrightarrow a \vee \neg a \qquad (1.4)$$

By performing this naming-operation on these statements, one is obviously changing their truth value. Even though we have one consistent, one inconsistent and one tautological statement on the left, all the statements become consistent after they have been named. This is because we in all the cases above can find a truth value for x_i that matches the one of the corresponding statement. This will not be the case for paradoxes, so our new named statements will be consistent if and only if they are not paradoxical.

Consider the liar sentence. It can be written as a named statement in the following way:

$$x \leftrightarrow \neg x \qquad (1.5)$$

This statement is obviously inconsistent, making it a paradox by our newly acquired definition. The study of *discourses* takes this formalization a step further.

1.2 Discourses

A propositional theory is in **graph normal form (GNF)** if all its formulae have the following form:

$$x \leftrightarrow \bigwedge_{i \in I_x} \neg y_i \qquad (1.6)$$

such that every variable occurs exactly once on the left of \leftrightarrow across all the formulae in the theory.

There is a simple translation from conjunctive normal form to graph normal form (shown in the appendix), showing that any propositional theory has an equisatisfiable GNF theory. By interpreting the variable on the left as the name of the statement on the right, like shown earlier, one can start using GNF to model meta-statements.

We now formally define a **discourse** to be a theory in GNF, and a **paradox** to simply be an inconsistent discourse.

¹**TODO:** bad wording?

We will later show a very handy correspondence between these discourse theories and certain graphs. The correspondence lets us not only decide the satisfiability of a discourse theory (i.e. whether or not it is paradoxical) by looking at certain properties of the corresponding graph. The properties in the graph also provide us with the satisfying models, if they exist. In order to express this logic/graph correspondence, we first need to establish some graph terminology.

1.3 Graphs, Kernels and Solutions

A directed graph (digraph) is a pair $\mathbf{G} = \langle G, N \rangle$ where G is a set of vertices while $N \subseteq G \times G$ is a binary relation representing the edges in \mathbf{G} . We use the notation $N(x)$ to denote the set of all vertices that are targeted by edges originating in x (successors of x). Similarly, $N^-(x)$ denotes the set of all vertices with edges targeting x (predecessors of x). We define these two predicates formally as follows:

$$N(x) := \{y \mid (x, y) \in N\} \quad (1.7)$$

$$N^-(x) := \{y \mid (y, x) \in N\} \quad (1.8)$$

A simple *path* is a sequence of distinct vertices x_1, x_2, \dots, x_n such that for any consecutive pair x_i, x_{i+1} from the sequence, we have $(x_i, x_{i+1}) \in N$. We say that two paths are *disjoint* if they do not share any vertices (possibly with the exception of their initial nodes).

The predicates N and N^- can be extended pointwise to sets in the following way:

$$N(X) = \bigcup_{x \in X} N(x) \quad (1.9)$$

$$N^-(X) = \bigcup_{x \in X} N^-(x) \quad (1.10)$$

A kernel is a set of vertices $K \subseteq G$ such that:

$$G \setminus K = N^-(K) \quad (1.11)$$

The above equivalence can be split up into two inclusions to be more easily understood:

$G \setminus K \subseteq N^-(K)$, saying that each vertex outside the kernel has to have an edge into the kernel (K is absorbing).

$N^-(K) \subseteq G \setminus K$, saying that each edge targeting a vertex within the kernel has to come from outside, thus no two vertices in the kernel are connected by an edge (K is independent).

Kernels have been of great interest over several decades, mainly within the fields of game theory and economics. The concept was first defined and used by Neumann and Morgenstern in [2]. In a graph representing some sort of a turn-based game, where vertices are states and edges are transitions, one can often work out winning strategies whenever one finds a kernel in the graph. Whenever one is outside of the kernel, one always has the possibility of moving inside the kernel (since the kernel is absorbing), while inside the kernel one *has* to move out of it (since the kernel is independent). If you are the player with the choice outside the kernel, you can control the game and choose to stabilize it by always moving inside the kernel, forcing the opponent to move out again.

Deciding the existence of kernels in finite graphs has been shown to be an NP-complete problem[3]. This should not be surprising, since we are in the middle of showing the equivalence between this problem and the problem of finding satisfying models of PL theories (SAT), which we know is NP-complete ².

We will get the correspondence between satisfying models of a discourse theory and kernels in a graph through an alternative, equivalent kernel definition called a *solution*.

Given a directed graph $\mathbf{G} = \langle G, N \rangle$, an assignment $\alpha \in 2^G$ is a function mapping every vertex in the graph to either 0 or 1. A **solution** is an assignment α such that for all $x \in G$:

$$\alpha(x) = 1 \iff \alpha(N(x)) = \{0\} \quad (1.12)$$

This means that for any node x , if x is assigned 1, then all its successors has to be assigned 0, and if x is assigned 0, then there has to exist a node assigned 1 among its successors. A consequence of this definition is that all sink nodes (nodes with no outgoing edges) in the graph have to be assigned 1, since it vacuously does not point to any node assigned 1. We use the notation $sol(\mathbf{G})$ to denote the set of all solutions of the graph \mathbf{G} .

1.4 Discourse Theories and Digraphs

As mentioned earlier, there is a close connection between (1) models of a discourse, (2) kernels of a graph and (3) solutions of a graph. While we have the equivalence between (2) and (3), we will now look at two functions connecting (1) and (2). This correspondence was shown by Roy T. Cook in [4]. We get the following definitions from [5]

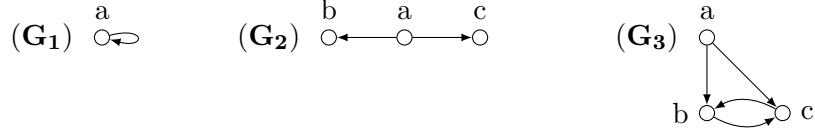
\mathcal{T} : translating a digraph \mathbf{G} into a corresponding theory $\mathcal{T}(\mathbf{G})$ such that $sol(\mathbf{G}) = mod(\mathcal{T}(\mathbf{G}))$.

²We are concerned with SAT over infinitary formulae in this paper, not the finite version from computer science.

\mathcal{G} : translating a theory T into a corresponding digraph $\mathcal{G}(T)$ such that $\text{mod}(T) = \text{sol}(\mathcal{G}(T))$.

Given any digraph \mathbf{G} we get the theory $\mathcal{T}(\mathbf{G})$ by taking, for each $x \in G$, the formula $x \leftrightarrow \bigwedge_{y \in N(x)} \neg y$ where $\bigwedge \emptyset = 1$.

Example 1.



Using the graphs from above, we get the following theories using \mathcal{T} :

$$\mathcal{T}(\mathbf{G}_1) = \{a \leftrightarrow \neg a\} \quad (1.13)$$

$$\mathcal{T}(\mathbf{G}_2) = \{a \leftrightarrow (\neg b \wedge \neg c), b, c\} \quad (1.14)$$

$$\mathcal{T}(\mathbf{G}_3) = \{a \leftrightarrow (\neg b \wedge \neg c), b \leftrightarrow \neg c, c \leftrightarrow \neg b\} \quad (1.15)$$

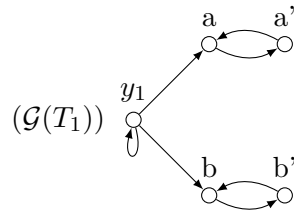
The fact that $\text{sol}(\mathbf{G}) = \text{mod}(\mathcal{T}(\mathbf{G}))$ is shown in [5]. Although not proving it, we can observe that \mathbf{G}_1 has no solution, just like its corresponding theory $\mathcal{T}(\mathbf{G}_1)$ has no models. \mathbf{G}_2 has one solution, where one assigns $a = 0, b = 1, c = 1$. This assignment also works as a model for $\mathcal{T}(\mathbf{G}_2)$. In \mathbf{G}_3 , we get two solutions, both with a assigned to 0, but with 0 and 1 distributed on b and c . These are also the only two models of $\mathcal{T}(\mathbf{G}_3)$.

Conversely, given any discourse theory T (in fact, this will work given any PL theory, since we can translate CNF to GNF), we can derive the corresponding graph $\mathcal{G}(T)$ in the following way: All variables in the theory are vertices, and for each formula $x \leftrightarrow \bigwedge_{i \in I_x} y_i$ make a directed edge $\langle x, y_i \rangle$ for each $i \in I_x$.

Example 2.

$$(a \leftrightarrow \neg a'), (a' \leftrightarrow \neg a), (b \leftrightarrow \neg b'), (b' \leftrightarrow \neg b), (y_1 \leftrightarrow (\neg a \wedge \neg b \wedge \neg y_1)) \quad (1.16)$$

Using \mathcal{G} on the above GNF theory gives us the following graph:



Again, will we not be proving the correspondence, but notice that T_1 has one model satisfying it, where $a = 0$. $\mathcal{G}(T_1)$ also has one solution, namely where $a = 0, a' = 1, y_1 =$

0. T_2 has three solutions, where either a , b or both are assigned 1. This reflects onto the graph since y_1 has to be assigned 0, thus forcing a or b to be assigned 1. The fact that \mathcal{G} gives us the correspondence we are looking for is shown in [5].

With the problem of solutions in the graph being equivalent with SAT, we get our final equivalence between kernels in the graph and satisfying models of the theory. This equivalence connects the fields of logic with the fields of graph theory: A graph has a kernel if and only if its corresponding discourse is consistent (non-paradoxical). A theory is paradoxical if and only if its corresponding graph has no kernels. Because of this tight link, we will often refer to graphs without kernels as paradoxical graphs.

The applicability of kernels should by now be obvious. In the next section we will review some of the various findings within Kernel Theory, and especially the findings related to infinitary graphs.

1.5 Results in Kernel Theory

Ultimately, the end goal of Kernel Theory would be to have an easy way to answer the question "Does this digraph have a kernel?" no matter the graph, and no matter the answer. We are not quite there, but a lot of work has been put into trying to identify special circumstances under which one is guaranteed to have (or guaranteed to not have) a kernel in the given graph. One of the results is the Richardson's Theorem, worked out by Moses Richardson in 1953:

Theorem 3. *If D is a finitary³ digraph without odd cycles, then D has a kernel[6].*

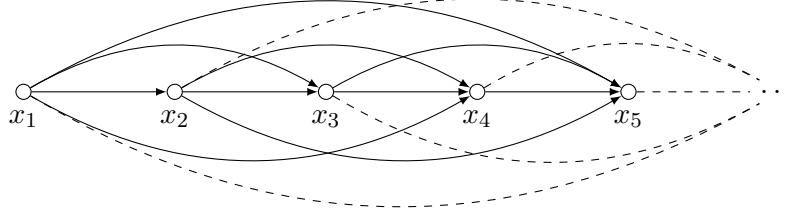
This theorem confirms that whenever dealing with finitary dags, for instance, one can be certain that its corresponding theory is consistent.

Intuitively, one might be tempted to believe that *all* digraphs without odd cycles have kernels, but this is not the case. Until now, our paradoxes have always been statements that – directly or indirectly – have been referring back to themselves (giving cycles in the graph) and thus causing a logical conflict, and it is hard to imagine any other way to construct paradoxical statements. The following construction will however reveal our lack of imagination.

The **Yablo Graph** is an example of an acyclic graph with no kernel[7]. It is constructed with an infinite set of vertices $\{x_i | i \in \mathbb{N}\}$ and a set of edges N such that $\langle x_i, x_j \rangle \in N$ iff

³In a finitary graph, every vertex has a finite number of out-neighbors; the graph has finite branching.

$i < j$.



Since there exist no two numbers $x, y \in \mathbb{N}$ such that $x < y$ and $y < x$, we get that the Yablo graph indeed is acyclic⁴. Furthermore, since any natural number has infinitely many numbers strictly larger than it, we get that all the vertices are infinitely branching, making the Yablo graph infinitary (not finitary).

The corresponding discourse theory of the Yablo graph would – informally – be the situation with an infinite number of statements, all saying “Every statement after this statement is false”.

We will later show formally that the Yablo-graph is indeed without a kernel, but for now the following explanation will do.

Let us assume that the Yablo-graph *has* a kernel and that the vertex x_a is in it. Then all the vertices to the right of x_a are necessarily outside of the kernel, including x_{a+1} . But if x_{a+1} is outside of the kernel, it has to point to a vertex on the inside. This is now impossible, since the out-neighborhood of x_{a+1} is a subset of the out-neighborhood of x_a . Since x_a was chosen without any restrictions, no vertex can be inside the kernel, making it empty. This is obviously not possible, so the Yablo-graph is without a kernel.

One thing should be mentioned at this point; neither odd cycles nor infinitely branching vertices *entail* that their respective graphs are paradoxical. The two following graphs illustrate this point:



The above graph contains an odd cycle, but the singleton set $\{x_2\}$ is a kernel.

The above graph has an infinitely branching vertex x_0 , but the infinite set $\{x_i \mid x > 0\}$ is a kernel.

It is shown in [5] that every digraph (with at least one edge) can be transformed into a infinitary dag⁵ such that α is a solution to the created dag if and only if it is a solution

⁴**TODO:** This is a bit over-simplified

⁵Directed acyclic graph

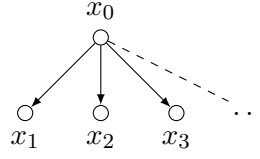


Figure 1.1

to the original digraph. This means that for any finitary graph that is paradoxical by the virtue of having an odd cycle, there is an infinitary, *acyclic* digraph that is also paradoxical. So, if one is trying to find ways to identify paradoxical graphs, one does only need to look at dags.

This result will be of great importance to us, enabling us to narrow our search space when looking for paradoxes.

1.6 Recognizing dags without kernels

Knowing that any graph can be translated to an equisatisfiable dag, the challenge is now to find sufficient conditions for dags to have kernels, even weaker than the one proved by Richardson (the fact that any finitary dag has a kernel is a direct consequence of Richardson's Theorem).

Michał Walicki has proposed the following thesis:

“If a dag has no kernel then it has a ray with infinitely many vertices dominating it.”

Some terminology (given a graph $\mathbf{G} = \langle G, N \rangle$): A *ray* is an semi-infinite path, i.e. an infinite sequence (x_1, x_2, \dots) of distinct vertices of G such that $(x_i, x_{i+1}) \in N$ for each i .

A vertex x_0 *dominates* a set of vertices $Y \subseteq G$ if there exists an infinite number of disjoint paths from x_0 to distinct vertices of Y .

The contrapositive of Walicki's thesis suggests a weaker condition for a kernel, since a dag having a ray with infinitely many vertices dominating it implies that the dag is infinitary.

1.7 Resolving GNF-theories

In this section, we will be presenting an inference system introduced by Walicki in [] which handles clausal theories induced from GNF-theories.

Recall that a theory written in GNF has formulae of the following form:

$$x \leftrightarrow \bigwedge_{i \in I_x} \neg y_i \quad (1.18)$$

Using simple operations only, one can manipulate these formulae into an equivalent set of clauses. We start by writing the above bi-implication as two implications:

$$x \rightarrow \bigwedge_{i \in I_x} \neg y_i \quad \text{and} \quad x \leftarrow \bigwedge_{i \in I_x} \neg y_i \quad (1.19)$$

The first implication can be rewritten in the following way:

$$x \rightarrow \bigwedge_{i \in I_x} \neg y_i = \neg x \vee \bigwedge_{i \in I_x} \neg y_i = \bigwedge_{i \in I_x} (\neg x \vee \neg y_i) = \bigwedge_{i \in I_x} \neg(x \wedge y_i) \quad (1.20)$$

The second implication can be rewritten in the following way:

$$x \leftarrow \bigwedge_{i \in I_x} \neg y_i = x \vee \neg \left(\bigwedge_{i \in I_x} \neg y_i \right) = x \vee \bigvee_{i \in I_x} y_i \quad (1.21)$$

By splitting the conjunction from the first implication up into individual clauses, we get the following two kinds of clauses for every variable x in the GNF theory:

$$\text{OR-clause: } x \vee \bigvee_{i \in I_x} y_i \quad (1.22)$$

$$\text{NAND-clauses: } \neg(x \wedge y_i), \text{ for every } i \in I_x \quad (1.23)$$

We will treat both the OR-clauses and the NAND-clauses as sets of atoms, denoting NAND-clauses $\neg(x \wedge y)$ as \overline{xy} and OR-clauses $x \vee y_1 \vee y_2 \vee y_3$ as $xy_1y_2y_3$. This enables us to state things like $\overline{xy} \subseteq \overline{xyz}$. A theory will – as expected – be a set of clauses.

If we interpret the initial GNF-theory as a graph $\mathbf{G} = \langle G, N \rangle$, for every vertex $x \in G$, there will be one OR-clause $\{x\} \cup N(x)$ and for every edge $\langle x, y \rangle \in N$ there will be a NAND-clause \overline{xy} . The graphs from Example 1 will have the following clausal theories:

$$\mathcal{T}(\mathbf{G}_1) = \{a, \overline{a}\} \quad (1.24)$$

$$\mathcal{T}(\mathbf{G}_2) = \{abc, b, c, \overline{ab}, \overline{ac}\} \quad (1.25)$$

$$\mathcal{T}(\mathbf{G}_3) = \{abc, bc, \overline{ab}, \overline{bc}\} \quad (1.26)$$

Further notation: $A \subseteq G$ denotes an OR-clause while $\overline{A} \subseteq G$ denotes a NAND-clause. Given a graph $\mathbf{G} = \langle G, N \rangle$, we denote the set of all NAND-clauses induced from the graph as NAND and all induced OR-clauses as OR . The combined set $\Gamma = \text{NAND} + \text{OR}$ will be our initial clauses in the inference system.

1.7.1 The inference system

We consider the following inference system, but we will focus mainly on proofs using the Axioms together with the (Rneg) rule.

$$(Ax) \quad \Gamma \vdash C, \quad \text{for } C \in \Gamma \quad (1.27)$$

$$(Rneg) \quad \frac{\{\Gamma \vdash \overline{a_i A_i} \mid i \in I\} \quad \Gamma \vdash \{a_i \mid i \in I\}}{\Gamma \vdash \bigcup_{i \in I} A_i} \quad (1.28)$$

$$(Rpos) \quad \frac{\Gamma \vdash A \quad \{\Gamma \vdash B_i K_i \mid i \in I\} \quad \{\Gamma \vdash \overline{a_i k} \mid i \in I, k \in K_i\}}{\Gamma \vdash (A \setminus \{a_i \mid i \in I\}) \cup \bigcup_{i \in I} B_i} \quad (1.29)$$

(Rneg) is creating NAND-clauses from NAND-clauses using OR as a side-condition. (Rpos) is creating OR-clauses from OR-clauses using NAND as a side-condition. In (Rneg), $\overline{a_i A_i}$ denotes the NAND $\{a_i\} \cup A_i$ with a potentially empty A_i .

The premise of the (Rneg) rule is a set of I NAND-clauses together with one OR-clause with I elements such that each atom a_i in the OR-clause is contained within a NAND-clause, and such that each NAND-clause contains an atom from the OR-clause. The correspondence between the NAND-clauses and the elements of the OR-clause should in other words be bijective. The conclusion is the union of all the NAND-clauses without their corresponding atom from the OR-clause.

Here are some examples of incorrect applications of the (Rneg)-rules, followed by some correct applications:

$$(1) \quad \frac{\overline{ax} \quad \overline{by} \quad \overline{cz}}{\overline{xyz}} abx \quad (3) \quad \frac{\overline{ax} \quad \overline{by}}{\overline{xy}} abx \quad (2) \quad \frac{\overline{ax} \quad \overline{by} \quad \overline{bz}}{\overline{xyz}} abx \quad (1.30)$$

(1) is incorrect because the NAND \overline{cz} contains no atoms from the OR abx . (2) is incorrect because the number of NAND-clauses does not match the length of the OR-clause. (3) is incorrect because there exist no bijective correspondence of the type described above.

$$(4) \quad \frac{\overline{ax} \quad \overline{by} \quad \overline{cz}}{\overline{xyz}} abc \quad (5) \quad \frac{\overline{ax} \quad \overline{b}}{\overline{x}} ab \quad (6) \quad \frac{\overline{ax} \quad \overline{by} \quad \overline{xyz}}{\overline{xyz}} abx \quad (1.31)$$

The three above applications are all correct, since all the atoms in each OR-clause get matched to exactly one NAND-clause in such a way that no NAND-clause stays unmatched.

We set no restrictions on the number and cardinality of our clauses, meaning that there might be an infinite number of clauses, and both the OR-clauses and the NAND-clauses might be either finite or infinite in size. Note that an infinite graph gives infinitely many NAND-clauses, while an infinitary graph also gives us infinitely long OR-clauses.

We study the refutation system that arises from the Axioms and the (Rneg)-rule, calling it Neg. It is shown in the submitted paper [michal-completeness] that Neg is sound

and refutationally complete for theories with only a countable number of OR-clauses. Soundness gives us that proving \overline{C} for any $C \subseteq G$ implies that the vertices in C cannot all be assigned 1 in the graph model. Refutational completeness gives us that whenever a graph/theory is inconsistent, we are able to prove \emptyset in Neg.

1.7.2 Inconsistency of the Yablo-graph

The inconsistency of the Yablo-graph is easily provable using Neg only. Since every vertex x_i (using the notation from earlier) has an edge to each vertex x_j where $j > i$, we get that every pair of distinct vertices is connected by an edge. This means that our set of axioms from the Yablo-graph looks like this:

$$\text{NAND} = \{\overline{x_i x_j} \mid i < j\} \quad \text{OR} = \{x_i x_{i+1} x_{i+2} \dots \mid i \in \mathbb{N}\} \quad (1.32)$$

For any vertex x_i from the Yablo-graph, we are now able to prove $\overline{x_i}$ in the following way:

$$\frac{\overline{x_i x_{i+1}} \quad \overline{x_i x_{i+2}} \quad \overline{x_i x_{i+3}} \quad \dots}{\overline{x_i}} \quad x_i x_{i+1} x_{i+2} \dots$$

Proving \emptyset is now simple:

$$\frac{\frac{\dots}{x_1} \quad \frac{\dots}{x_2} \quad \frac{\dots}{x_3} \quad \dots}{\emptyset} \quad x_1 x_2 x_3 \dots$$

A less trivial inconsistency proof is the one on the *Stretched Yablo-graph*. This proof can be found in the appendix together with the definition of Stretched Yablo.

It is worth mentioning that even though our focus has been – and will be – on theories originating from graphs, the results on soundness and completeness holds for any theory consisting of a set of NANDs and a set of ORs.

An example of this is the pigeonhole problem which easily can be represented as a set of NANDs and ORs, but does not directly correspond to a graph (it can of course be translated to a graph theory, like any other propositional theory). Proofs in Neg of pigeonhole problems of different size are to be found in the appendix⁶.

1.8 This thesis

1.8.1 Motivation

Having that (Neg) is both sound and refutationally complete, results linking inconsistency proofs in (Neg) to structures in graphs is of great interest to us, since it can

⁶**TODO:** some reference system needed

potentially further weaken the current conditions we have for kernels in graphs. This thesis will thus be concerned mainly with two kinds of questions;

1. Are there any characterizing qualities of the proof system (Neg), potentially under certain restrictions?
2. If the axioms come from a graph, are there any provable results in (Neg) that correspond to certain structures in the graph?

The first question covers qualities like proof length (complexity), size and number of clauses used throughout a proof, the existence of certain clauses in certain proofs, or maybe even the existence of an entire proof normal form. For instance, an answer to the following question could be of great interest to us: “Are there certain kinds of clauses that needs to be present in a proof in order to prove \emptyset ?”.

The second question covers questions like “What does it mean for the graph that its axioms can prove \emptyset ?”. Of course, we know the answer to this question, it means that the graph has no kernel (by soundness of (Neg)). But what about the provability of \bar{x} for a node x in the graph? What does that mean for the graph and the node x ?

Finding a type of clause neccesarly present in an inconsistency proof, together with results on the graph-structural counterpart of that type of clause, gives us in combination a graph-structure neccesarly present in a kernel-free graph (by soundness of (Neg)).

1.8.2 Thesis Overview

THIS SECTION IS NOT UP TO DATE WITH THE FOLLOWING CHAPTERS Section 2 will take a look at the potential to control the size of NAND-clauses in an inconsistency proof without any restrictions on the axioms. Section 3 will study the same potential as above, but with graph-based axioms. We are in both cases above particularly concerned with whether we are able to restrict our proofs to only use binary NAND-clauses. Section 4 will explore the graph structure that corresponds to unary and binary NAND-clauses, and explain their potential importance.

Chapter 2

NAND-clauses in graphs

In this chapter we will motivate the search for graph structural equivalents of unary and binary NAND-clauses proven in Neg. An actual graph structure corresponding to the NAND-clauses will not be presented, but some necessary conditions will be shown for both directions. In other words, we will search for graph structures implying certain provable NAND-clauses, as well as what certain provable NAND-clauses implies about the corresponding graph.

2.1 Unary and binary NAND-clauses

Since our proof system, Neg, only has one rule, the last step of an inconsistency proof will always look the same:

$$\frac{\overline{x_1} \quad \overline{x_2} \quad \overline{x_3} \quad \dots}{\emptyset} x_1 x_2 x_3 \dots$$

The premise will always consist of a collection of NAND-clauses, each of length 1, together with an OR-clause equal to the union of all the NAND-clauses. It is easy to see that none of the NAND-clauses can be larger than unary, since that would result in a non-empty NAND-clause in the conclusion. The OR-clause has to equal the union of the NAND-clauses simply by definition of the (RNeg)-rule.

This fact was also observed and formalized by Walicki[**michal-completeness**]:

$$\Gamma \vdash_{Neg} \{\} \Leftrightarrow \exists K \in \text{OR} : (\forall k \in K : \Gamma \vdash_{Neg} \overline{k}) \quad (2.1)$$

We know from the definition that any OR-clause used in the proof system corresponds to a single vertex with its successors in the graph. We do however not know what the NAND-clauses of length 1 might correspond to. Knowing this would, by soundness and

completeness of Neg, give us a complete picture of the graph structure needed for a kernel not to exist.¹

The only thing we *do* know about unary NAND-clauses is that they correspond to vertices that are assigned 0 in all models of the graph. We get this from soundness of Neg. This is however not the graph structural property we are ultimately looking for, but we can at least say that we have reduced the question “What does an inconsistent graph look like?” to the question “What does a provably false vertex look like?”.

So what does a unary NAND-clause proven in Neg correspond to in the graph? Similarly to a proof of \emptyset , there is really just one way to prove a unary NAND-clause:

$$\frac{\overline{xy_1} \quad \overline{xy_2} \quad \dots \quad \overline{y_n} \quad \overline{y_{n+1}} \quad \dots}{\overline{x}} \quad y_1 y_2 \dots y_n y_{n+1} \dots$$

Any derivation of a unary NAND-clause \overline{x} must end with a rule application using $K \in \text{OR}$ where the NAND-clauses in the premises are either unary, \overline{k} or binary, \overline{xk} , both such that $k \in K$.

We formalize this observation in the following way:

$$\Gamma \vdash_{\text{Neg}} \overline{x} \Leftrightarrow \exists K \in \text{OR} : (\forall k \in K : \Gamma \vdash_{\text{Neg}} \overline{kx} \vee \Gamma \vdash_{\text{Neg}} \overline{k}) \wedge (\exists k \in K : \Gamma \vdash_{\text{Neg}} \overline{kx}) \quad (2.2)$$

Just as we reduced the problem of inconsistency to the problem of unary NAND-clauses, we are able to reduce the problem further to binary NAND-clauses. We could even continue the reduction further to ternary clauses, quaternary clauses and so on, but without a change of strategy at some point, this seems pointless.

Our current reduction lets us ask how two vertices x, y are *connected* in the graph when their binary NAND \overline{xy} is proven in Neg. Observe that we have parts of this correspondence down already, with our axioms being all binary. Vertices directly connected by an edge must therefore be a part of the corresponding graph structure.

Finding some graph structural relation that corresponds to binary NAND-clauses might give us some knowledge about a potential standard form of binary-NAND-proofs.

¹**TODO:** Currently brushing over the restrictions set to the proofs of completeness and soundness. Will fix this later.

Chapter 3

Expressional power of binary NAND-clauses

In this chapter we will investigate the statement that whenever a theory is inconsistent, it can be proven in Neg using unary and binary NAND-clauses only. Such a result would be significant both because it would be a strong property for a proof system to have in general, but also because it could potentially help us to further characterize a kernel-free graph

Having that any inconsistency can be proven in Neg using unary and binary NAND-clauses only, might imply a similar property in kernel-free graphs, namely that inconsistencies can be described as a collection of pairwise structural relations between vertices.

We will check the statement first for non-graph-based theories and then for theories based on graphs with different properties.

3.1 Proving inconsistency of unrestricted theories

This section will show that there exist inconsistencies in non-graph-based theories that are not possible to prove in Neg using only unary and binary NAND-clauses. We will prove the pigeonhole principle to exemplify this.

The pigeonhole principle states that whenever you have n pigeons and m holes such that $n > m$, then at least one hole must contain more than one pigeon. We can prove this principle in Neg by proving an inconsistency in the theory that (1) all n pigeons are contained in a hole and (2) each hole contains only one pigeon.

Letting the atom x_i denote the pigeon i occupying the hole x , the above theory, with n

pigeons and m holes, can be formalized in the following way:

$$(1): \{ 1_i 2_i 3_i \dots m_i \mid i \leq n \} \quad (3.1)$$

$$(2): \{ x_i x_j \mid i < j \leq n, x \leq m \} \quad (3.2)$$

3.1.1 Proving the pigeonhole principle - 4 pigeons, 3 holes

The pigeonhole principle for 4 pigeons and 3 holes will have the following axioms:

$$\text{OR} = \{ 1_1 2_1 3_1, 1_2 2_2 3_2, 1_3 2_3 3_3, 1_4 2_4 3_4 \} \quad (3.3)$$

$$\text{NAND} = \left\{ \begin{array}{c} \overline{1_1 1_2}, \overline{1_1 1_3}, \overline{1_1 1_4}, \overline{1_2 1_3}, \overline{1_2 1_4}, \overline{1_3 1_4}, \\ \overline{2_1 2_2}, \overline{2_1 2_3}, \overline{2_1 2_4}, \overline{2_2 2_3}, \overline{2_2 2_4}, \overline{2_3 2_4}, \\ \overline{3_1 2_2}, \overline{3_1 3_3}, \overline{3_1 3_4}, \overline{3_2 3_3}, \overline{3_2 3_4}, \overline{3_3 3_4} \end{array} \right\} \quad (3.4)$$

From the axioms above we can prove inconsistency in the following way:

Proving $\overline{1_4}$:

$$\frac{\frac{\overline{1_3 1_4} \quad \overline{1_2 1_4} \quad \overline{2_2 2_3} \quad \overline{1_1 1_4} \quad \overline{2_1 2_3} \quad \overline{3_1 3_2}}{\overline{2_3 1_4} \quad \overline{3_2 2_3 1_4}} \quad 1_1 2_1 3_1 \quad \overline{1_2 1_4} \quad \overline{1_1 1_4} \quad \overline{2_1 2_2} \quad \overline{3_1 3_3} \quad 1_1 2_1 3_1 \quad \overline{3_2 3_3}}{\overline{3_3 1_4} \quad 1_3 2_3 3_3} \quad 1_2 2_2 3_2$$

$$\frac{\overline{1_3 1_4} \quad \overline{2_3 1_4}}{\overline{1_4}}$$

Proving $\overline{2_4}$:

$$\frac{\frac{\overline{1_2 1_3} \quad \overline{2_2 2_4} \quad \overline{1_1 1_3} \quad \overline{2_1 2_4} \quad \overline{3_1 3_2}}{\overline{3_2 1_3 2_4}} \quad 1_1 2_1 3_1 \quad 1_1 2_1 3_1 \quad \overline{1_1 1_2} \quad \overline{2_1 2_4} \quad \overline{3_1 3_3}}{\overline{1_3 2_4} \quad \overline{2_3 2_4} \quad \overline{3_3 2_4}} \quad 1_2 2_2 3_2 \quad \overline{1_2 2_3 2_4} \quad \overline{2_2 2_4} \quad \overline{3_2 3_3}}{\overline{2_4}} \quad 1_3 2_3 3_3$$

Proving $\overline{3_4}$:

$$\frac{\frac{1_2 2_2 3_2 \quad \overline{1_2 1_3} \quad 1_1 2_1 3_1 \quad \overline{1_1 1_3} \quad \overline{2_1 2_2} \quad \overline{3_1 3_4}}{\overline{2_2 1_3 3_4}} \quad \overline{3_2 3_4} \quad 1_1 2_1 3_1 \quad \overline{1_1 1_2} \quad \overline{2_1 2_3} \quad \overline{3_1 3_4}}{\overline{1_3 3_4} \quad 1_2 2_2 3_2 \quad \overline{1_2 2_3 3_4} \quad \overline{2_2 2_3} \quad \overline{3_2 3_4}} \quad 1_3 2_3 3_3 \quad \overline{1_3 3_4} \quad \overline{2_3 3_4} \quad \overline{3_3 3_4}}{\overline{3_4}}$$

With these three unary clauses proven, we can now prove \emptyset in one step:

$$1_4 2_4 3_4 \quad \frac{\overline{1_4} \quad \overline{2_4} \quad \overline{3_4}}{\emptyset}$$

Chapter 4

Proofs

4.1 Translating CNF to GNF

Since any PL theory can be expressed in CNF, showing that any theory P in CNF can be translated to a theory R in GNF such that P and R are equisatisfiable gives us that any PL theory has an equisatisfiable GNF.

Given any CNF theory P , for each formula in it, follow the steps below to acquire its corresponding GNF theory.

Step 1: For each literal $\neg x_i$ in the formula, introduce a fresh variable x'_i and add the following two GNF formulae to P : $x'_i \leftrightarrow \neg x_i, x_i \leftrightarrow \neg x'_i$, (unless this has already been done while translating an earlier formula in the theory).

Step 2: In each clause of the formula, replace every negative literal $\neg x_i$ with its corresponding x'_i from step 1. Every clause in the formula does now contain positive literals only.

Step 3: For every clause $(x_1 \vee x_2 \vee \dots \vee x_n)$, replace it with the following formula, where y is fresh:

$$y \leftrightarrow (\neg x_1 \wedge \neg x_2 \wedge \dots \wedge \neg x_n \wedge \neg y)$$

This formula is equisatisfiable with our original clause.

The combined set of all these acquired formulae will make up our GNF-theory. We have only added proper GNF formulae and all variables appear to the left in exactly one clause, so R will indeed be a GNF theory.

Step 1 is adding a bi-implication formula that is always satisfiable, since one of two variables in it is fresh. One can think of it as a labelling of an already existing variable. Thus, adding these formulae to a theory does not change its satisfiability/non-satisfiability.

Step 2 is replacing each negative literal with its label, also not changing the satisfiability.

Step 3 is replacing one formula with an equisatisfiable formula, thus not changing the overall satisfiability.

Since none of the steps above change the satisfiability of the theory, we can conclude that our acquired theory is equisatisfiable with to original CNF theory.

Below are some examples of CNF-theories with their corresponding GNF-theories.

Example 4.

$$\text{CNF: } (a \vee b) \tag{4.1}$$

$$\text{GNF: } (a \leftrightarrow \neg a'), (a' \leftrightarrow \neg a), (b \leftrightarrow \neg b'), (b' \leftrightarrow \neg b), (y_1 \leftrightarrow (\neg a \wedge \neg b \wedge \neg y_1)) \tag{4.2}$$

$$\tag{4.3}$$

$$\text{CNF: } (\neg a) \tag{4.4}$$

$$\text{GNF: } (a \leftrightarrow \neg a'), (a' \leftrightarrow \neg a), (y_1 \leftrightarrow (\neg a' \wedge \neg y_1)) \tag{4.5}$$

$$\tag{4.6}$$

4.2 Inconsistency of Stretched Yablo

One variation of the Yablo-graph is the Stretched Yablo-graph, presented by Walicki in [michal-completeness]. While the Yablo-graph is a ray where each vertex on it has direct edges to all the vertices after it, the Stretched Yablo-graph is a ray where each vertex has *disjoint paths* (with certain properties depending on the variation of Stretched Yablo) to all vertices in after it.

The variation on which we will be proving an inconsistency is one with a set of core vertices \mathbb{N} where each vertex x has a disjoint path to every vertex y after it, such that the length of the path is $(2 \times (y - x) - 1)$.

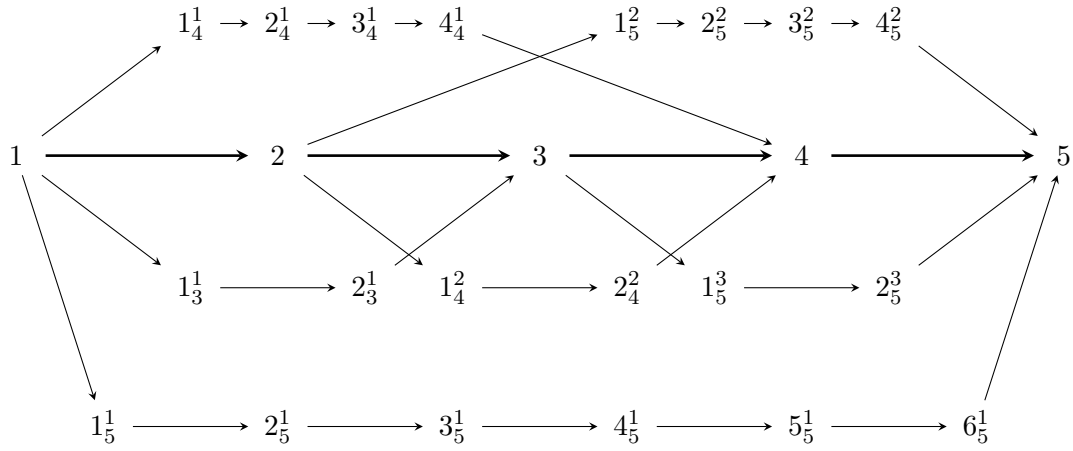


Figure 4.1: Stretched Yablo

Shown above are parts of the Stretched Yablo-graph described. We denote the core vertices by natural numbers. The peripheral vertices contained in each path is denoted by n_y^x where x and y is the source and target, respectively, of the path in which the vertex is contained. n denoted the relative position on the path.

Note that since two core vertices x and y have a peripheral path between them of length $L_y^x = 2(y - x) - 1$ (from the definition above), each path consists of $L_y^x - 1 = 2(y - x) - 2$ peripheral vertices. The vertex pointing back onto the core ray will thus always be the vertex $(2(y - x) - 2)_y^x$.

Based on the given definition of Stretched Yablo and the notation from above, we get

the following sets of axioms:

$$N1a = \{ \overline{x(x+1)} \mid x \in \mathbb{N} \} \quad (4.7)$$

$$N1b = \{ \overline{x1_y^x} \mid x, y \in \mathbb{N}, x+1 < y \} \quad (4.8)$$

$$N2 = \{ \overline{n_y^x(n+1)_y^x} \mid n, x, y \in \mathbb{N}, x+1 < y, n < 2(y-x) - 2 \} \quad (4.9)$$

$$N3 = \{ \overline{n_y^x y} \mid n, x, y \in \mathbb{N}, x+1 < y, n = 2(y-x) - 2 \} \quad (4.10)$$

$$O1 = \{ x(x+1)1_{x+2}^x 1_{x+3}^x \dots \mid x \in \mathbb{N} \} \quad (4.11)$$

$$O2 = \{ \overline{n_y^x(n+1)_y^x} \mid n, x, y \in \mathbb{N}, x+1 < y, n < 2(y-x) - 2 \} \quad (4.12)$$

$$O3 = \{ \overline{(n_y^x y)} \mid n, x, y \in \mathbb{N}, x+1 < y, n = 2(y-x) - 2 \} \quad (4.13)$$

$$\text{NAND} = N1a \cup N1b \cup N2 \cup N3 \quad \text{OR} = O1 \cup O2 \cup O3 \quad (4.14)$$

One can identify the different axioms by looking at the figure above¹: $N1a$ are the edges making up the core ray of the graph, $N1b$ are the edges going from a core vertex onto a peripheral vertex, $N2$ are the edges between two peripheral vertices and $N3$ are the edges going from a peripheral vertex back onto a core vertex. $O1$ are the vertices on the core ray, $O2$ and $O3$ are the peripheral vertices, $O3$ being the ones that points back to a core vertex.

In order to prove \emptyset from these axiom, we first prove three sets of intermediate clauses:

$$D1 = \{ \overline{xy} \mid x, y \in \mathbb{N}, x \neq y \} \quad (4.15)$$

$$D1b = \{ \overline{1_y^x n_y^x} \mid x, y \in \mathbb{N}, x+1 < y, n = 2(y-x) - 2 \} \quad (4.16)$$

$$D2 = \{ \overline{x1_z^y} \mid x, y, z \in \mathbb{N}, x \neq z, y+1 < z \} \quad (4.17)$$

$$D3 = \{ \overline{1_x^1 1_z^y} \mid x, y, z \in \mathbb{N}, x \neq z, y+1 < z \} \quad (4.18)$$

4.2.1 Proof, D1 and D1b

We already have \overline{xy} from our axioms whenever $y = x \pm 1$, so let us assume that either $y < x - 1$ or $x + 1 < y$. In both these cases, we get that there exists a peripheral path between x and y . Let us – without loss in generality – assume that $x + 1 < y$. We thus have a path from x to y containing $n = 2 \times (y - x) - 2$ peripheral vertices. Consider now the following proof:

¹recall how NANDs corresponds to edges, while ORs corresponds to vertices

$$\begin{array}{c}
\begin{array}{c}
(N1b) \quad (N2) \\
\frac{1_y^x 2_y^{2x}}{\frac{\overline{x 1_y^x} \quad \overline{2_y^{2x} 3_y^x}}{x 3_y^x}} \quad (N2) \\
(O2) \quad 3_y^x 4_y^x \quad \frac{\overline{4_y^x 5_y^x}}{x 5_y^x} \quad (N2) \\
(O2) \quad 5_y^x 6_y^x \quad \frac{\overline{6_y^x 7_y^x}}{x 7_y^x} \\
\vdots \\
(O2) \quad (n-1)_y^x n_y^x \quad \frac{\overline{x(n-1)_y^x} \quad \overline{n_y^x y}}{xy}
\end{array}
\end{array}$$

Intuitively, this proof follows along the peripheral path between the two core vertices. It is important to notice that since all our peripheral paths are odd in length, $(n-1)$ is always odd. This guarantees that we are able to prove $\overline{x(n-1)_y^x}$ in $n/2$ steps using the strategy above.

Using x, y, n from above, the proof of D1b follows a similar pattern as D1:

$$\begin{array}{c}
\begin{array}{c}
(N2) \quad (N2) \\
\frac{2_y^{2x} 3_y^x}{\frac{\overline{1_y^x 2_y^{2x}} \quad \overline{3_y^{2x} 4_y^x}}{1_y^x 4_y^x}} \quad (N2) \\
(O2) \quad 4_y^x 5_y^x \quad \frac{\overline{5_y^x 6_y^x}}{1_y^x 6_y^x} \quad (N2) \\
(O2) \quad 6_y^x 7_y^x \quad \frac{\overline{7_y^x 8_y^x}}{1_y^x 8_y^x} \\
\vdots \\
(O2) \quad (n-2)_y^x (n-1)_y^x \quad \frac{\overline{1_y^x (n-2)_y^x} \quad \overline{(n-1)_y^x n_y^x}}{1_y^x n_y^x}
\end{array}
\end{array}$$

Just like with D1, it is crucial that the path is odd in order for this proof strategy to work.

4.2.2 Proof, D2

Again, we have a trivial case where $x = y$ making the clause simply an instance of N1b. We therefore continue under the assumption that $x \neq y$, together with the other restrictions from the definition of D2. Using the same notation as above, where n is the number of vertices in the peripheral path between y and z , we now get the following short proof:

$$\begin{array}{c}
\begin{array}{c}
(D1) \quad (D1b) \\
(O3) \quad n_z^y z \quad \frac{\overline{xz} \quad \overline{1_z^y n_z^y}}{x 1_z^y}
\end{array}
\end{array}$$

Using $D1$ and $D1b$ in the above proof, restricts us to cases where $x \neq z$ and where $y + 1 < z$, but these are restrictions we have already assumed.

4.2.3 Proof, D3

n now denotes the number of vertices in the peripheral path between y and z .

$$(O3) \quad n_z^y z \quad \frac{\begin{array}{c} (D2) \\ \overline{1_x^1 z} \end{array} \quad \begin{array}{c} (D1b) \\ \overline{1_z^y n_z^y} \end{array}}{\overline{1_x^1 1_z^y}}$$

We are again restricting our cases by using $D3$ and $D1b$, but these restrictions are already assumed.

4.2.4 Proof, \emptyset

We will prove the inconsistency of Stretched Yablo by first proving $\overline{1}, \overline{2}, \overline{1_3^1}, \overline{1_4^1 1_5^1} \dots$.

Proving $\overline{1}$:

$$(O1) \quad 231_4^2 1_5^2 1_6^2 \dots \quad \frac{\begin{array}{c} (N1) \\ \overline{1_2} \end{array} \quad \begin{array}{c} (D1) \\ \overline{1_3} \end{array} \quad \begin{array}{c} (D2) \\ \overline{11_4^2} \end{array} \quad \begin{array}{c} (D2) \\ \overline{11_5^2} \end{array} \quad \begin{array}{c} (D2) \\ \overline{11_6^2} \end{array} \quad \dots}{\overline{1}}$$

Proving $\overline{2}$:

$$(O1) \quad 341_5^3 1_6^3 1_7^3 \dots \quad \frac{\begin{array}{c} (N1) \\ \overline{2_3} \end{array} \quad \begin{array}{c} (D1) \\ \overline{2_4} \end{array} \quad \begin{array}{c} (D2) \\ \overline{11_5^3} \end{array} \quad \begin{array}{c} (D2) \\ \overline{11_6^3} \end{array} \quad \begin{array}{c} (D2) \\ \overline{11_7^3} \end{array} \quad \dots}{\overline{2}}$$

Proving $\overline{1_3^1}$:

$$(O1) \quad 451_6^4 1_7^4 1_8^4 \dots \quad \frac{\begin{array}{c} (D2) \\ \overline{1_3^1 4} \end{array} \quad \begin{array}{c} (D2) \\ \overline{1_3^1 5} \end{array} \quad \begin{array}{c} (D3) \\ \overline{1_3^1 1_6^4} \end{array} \quad \begin{array}{c} (D3) \\ \overline{1_3^1 1_7^4} \end{array} \quad \begin{array}{c} (D3) \\ \overline{1_3^1 1_8^4} \end{array} \quad \dots}{\overline{1_3^1}}$$

Proving $\overline{1_4^1}$:

$$(O1) \quad 561_7^5 1_8^5 1_9^5 \dots \quad \frac{\begin{array}{c} (D2) \\ \overline{1_4^1 5} \end{array} \quad \begin{array}{c} (D2) \\ \overline{1_4^1 6} \end{array} \quad \begin{array}{c} (D3) \\ \overline{1_3^1 1_7^5} \end{array} \quad \begin{array}{c} (D3) \\ \overline{1_3^1 1_8^5} \end{array} \quad \begin{array}{c} (D3) \\ \overline{1_3^1 1_9^5} \end{array} \quad \dots}{\overline{1_4^1}}$$

The clauses $\overline{1_4^1}, \overline{1_5^1}, \overline{1_5^1}, \overline{1_7^1}, \dots$ can be proven in exactly the same manner as above. With the results from above we can finally prove \emptyset :

$$(O1) \quad 121_3^1 1_4^1 1_5^1 \dots \quad \frac{\overline{1} \quad \overline{2} \quad \overline{1_3^1} \quad \overline{1_4^1} \quad \overline{1_5^1} \quad \dots}{\emptyset}$$

Having proved \emptyset in Neg, we get from soundness of Neg that our Stretched Yablo-graph is indeed inconsistent.

Bibliography

- [1] G. Priest and F. Berto, “Dialetheism,” in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., Summer 2013, 2013. [Online]. Available: <http://plato.stanford.edu/archives/sum2013/entries/dialetheism/>.
- [2] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*. Princeton University Press, 1944 (1947).
- [3] V. Chvátal, “On the computational complexity of finding a kernel. technical report crm-300,” Centre de Recherches Mathématiques, Université de Montréal, Tech. Rep., 1973.
- [4] R. T. Cook, “Patterns of paradox,” *The Journal of Symbolic Logic*, vol. 69(3), pp. 767–774, 2004.
- [5] M. Bezem, C. Grabmeyer, and M. Walicki, “Expressive power of digraph solvability,” *Annals of Pure and Applied Logic*, vol. 163, pp. 200–213, 2012.
- [6] M. Richardson, “Solutions of irreflexive relations,” *The Annals of Mathematics, Second Series*, vol. 58(3), pp. 573–590, 1953.
- [7] S. Yablo, “Paradox without self-reference,” *Analysis*, vol. 53(4), pp. 251–252, 1993.