Diagnosing paradoxes

Michal Walicki
Department of Informatics
University of Bergen
michal@ii.uib.no

Sjur Dyrkolbotn
Department of Civil Engineering
Western Norway University of Applied Sciences
sdy@hvl.no

ABSTRACT

Graph normal form, GNF, [3], was used in [9, 12] for analysing paradoxes in finitary propositional discourses, with the semantics – equivalent to classical logic – defined by kernels of (di)graphs (equivalently, stable sets in finitary argumentation frameworks). Turning to the general case - infinitary graphs - this paper formulates a specific graph structure, underlying the liar and Yablo's paradox, which is conjectured to be present in all paradoxes (equivalently, all infinitary argumentation frameworks that do not admit stable sets). A theorem confirms this for graphs with finitely many ends and some with infinitely many ends, but the general case remains open. The paper then turns to proof theory, to deliver a new diagnostic tool that can be used to identify paradoxical discourses. The paper presents a simple system of infinitary, resolution-based reasoning with GNF theories, allowing us to determine the presence of a paradox in all countable discourses. The system is refutationally complete for the classical semantics, but when used for nonrefutational deduction it is not explosive and identifies in a paradoxical discourse a maximal consistent subdiscourse, including its classical consequences. The notion of a semikernel provides the basis for a novel paraconsistent semantics, which coincides with the classical semantics for consistent theories. It gives also a new semantics for argumentation based on admissible sets.

ACM Reference format:
Michal Walicki and Sjur Dyrkolbotn. 2017. Diagnosing paradoxes. In *Proceedings of TARK 2017, Liverpool, UK, July 2017 (TARK'17),* 8 pages.
DOI: 10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

An informal discourse, represented by just writing its statements in some logical language, can be analyzed for consistency or validity, but hardly for paradoxicality. Paradox does not amount to the inconsistency of the discourse but of its truth-theory, which means here, roughly, the collection of T-schemata for the discourse's statements, [9, 12, 18]. There is nothing paradoxical about $a \wedge \neg a$. Its propositional T-schema, $f \leftrightarrow (a \wedge \neg a)$, classifies simply this statement, called now f, as false. When there are no references between statements, the truth-theory becomes such a trivially satisfiable repetition of each statment in an equivalence to its unique identifier. (For instance, axiomatization of T-predicate by T-schemata applied only to arithmetic sentences gives a conservative extension of PA.)

One direction of research, on the formal theory of truth, tries to stretch this as far as possible, studying possible restrictions on the T-schemata, expressed in axiomatizations of T-predicate, preventing inconsistency. Another line of research looks instead for patterns of applications of T-schemata which lead to paradox. In essence, both do the same, investigating sufficient conditions for the consistency of a truth-theory. But the specific questions, techniques and results are different. While logic dominates the former, the latter increasingly uses graph theory, [2, 9, 12–14, 18], as we also do here.

Section 2 recalls, after [9, 12], the equivalence of propositional theories in GNF and digraphs, with models of a theory corresponding to kernels of its digraph. The link with formal argumentation theory is also discussed, along with some key semantic concepts, including the notion of a semikernel (known as an admissible set in argumentation theory). Section 3 presents the main conjecture, specifying a general graph pattern underlying – presumably – all paradoxes. New results confirm it for several special cases, in which exclusion of the pattern guarantees absence of paradox. However, the conjecture, even if true in its full generality, provides only necessary conditions of paradox. Any particular case, which does not fall under this general pattern, still needs an individual analysis. Section 4 presents therefore a resolution-based reasoning system from [18], which is capable of proving inconsistency of the truth-theory for any countable, paradoxical discourse. Although the reasoning utilizes classical resolution, its nonrefutational use leads to interesting paraconsistent features. Section 5 shows how, in the presence of inconsistency, the system proves only classical consequences of (an appropriately defined) maximal consistent subdiscourse. Section 6 provides a new semantics for the reasoning system, which is also a new semantics for argumentation based on admissible sets. Its key feature as an argumentation semantics is that it uniquely characterises the undecided arguments, while also agreeing with the stable semantics for all frameworks that admit stable sets.

2 BACKGROUND

A propositional formula is in graph normal form, GNF, when it has the form

$$x \leftrightarrow \bigwedge_{i \in I_x} \neg y_i, \tag{2.1}$$

where all x, y_i are atoms (propositional variables). When $I_x = \varnothing$, this is identified with x. A theory is in GNF when all formulae are in GNF and every atom occurs in such a formula exactly once unnegated, i.e., on the left of $\leftrightarrow$.[1] A *discourse* is a theory in GNF

[1]The formula $a \leftrightarrow \neg b$ is in GNF but the theory $\{a \leftrightarrow \neg b\}$ is not, due to the loose b. Such cases can be treated as abbreviations for GNF theories, here, with a fresh atom $\underline{b}$ and two additional formulae $b \leftrightarrow \neg\underline{b}$ and $\underline{b} \leftrightarrow \neg b$.

TARK'17, Liverpool, UK

DOI: 10.1145/nnnnnnn.nnnnnnn

and *paradox* is defined as an inconsistent discourse. Plausibility of this definition, implicit in [9], was argued and exemplified in [12], so we give only one illustration.

Example 2.2. *Let Θ_1 be the following discourse:*

a. This and the next statement are false.	$a \leftrightarrow \neg a \wedge \neg b$
b. The next statement is false.	$b \leftrightarrow \neg c$
c. The previous statement is false.	$c \leftrightarrow \neg b$

Making b true and a and c false, gives a model, so that Θ_1 does not involve any paradox. Adding the fourth statement:

d. This and the previous statement are false.	$d \leftrightarrow \neg d \wedge \neg c$

gives the discourse Θ_2, where paradox is unavoidable.

GNF is indeed a normal form: every theory in (infinitary) propositional logic $\mathcal{L}_\kappa$ has an equisatisfiable one in GNF, [3], so we assume theories to be in GNF without mentioning it.[2] The semantics is defined in the standard way and thus, although focusing on the paradoxical character of discourses, we indirectly address consistency in infinitary logic in general.

The standard semantics has an equivalent formulation in terms of graph kernels, [9, 12], which then enables a seamless transition between the classical and less classical logic. A graph (meaning in this paper "directed graph") is a pair $\mathsf{G} = \langle G, \mathsf{N} \rangle$ with $\mathsf{N} \subseteq G \times G$. We denote $\mathsf{N}(x) = \{y \in G \mid \mathsf{N}(x,y)\}$, $\mathsf{P}(x) = \{y \in G \mid x \in \mathsf{N}(y)\}$, and extend pointwise such notation to sets, i.e., $\mathsf{P}(X) = \bigcup_{x \in X} \mathsf{P}(x)$, etc. P^* denotes reflexive, transitive closure of P. A *kernel* of a graph G is a subset $K \subseteq G$ which is independent (no edges between vertices in K) and dominating (every vertex in $G \setminus K$ has an edge to some vertex in K), namely, such that $\mathsf{P}(K) = G \setminus K$. $Ker(\mathsf{G})$ denotes kernels of G.

We let $\bar{\mathsf{G}}$ denote the digraph obtained from G by reversing the direction of all edges, i.e., such that $\mathsf{N}_{\bar{\mathsf{G}}}(x) = \mathsf{P}(x)$ and $\mathsf{P}_{\bar{\mathsf{G}}}(x) = \mathsf{N}(x)$ for all $x \in G$. A kernel of G then corresponds to a so-called *stable set* in $\bar{\mathsf{G}}$ and vice versa [11]. Stable sets are studied in the theory of formal argumentation, where the link with kernels in graph theory is seldom noted.[3] However, these two notions are effectively the same and for lack of space we conflate the two in this paper. The reader interested in argumentation theory then has two choices: (i) $(x, y) \in \mathsf{N}$ can be read as encoding an "attack" made by y on x or (ii) the reader can replace every occurence of G by $\bar{\mathsf{G}}$, to formally translate between the two formalisms. Either way, what follows applies both to argumentation frameworks and propositional discourses. We recommend option (i), as we will speak of attacks in G based on this convention.

In addition to the notion of a kernel, we will use the notion of a semikernel, corresponding to the notion of an admissible set in argumentation theory [11]. Formally, a *semikernel* of G is a set $S \subseteq G$ such that

$$\mathsf{N}(S) \overbrace{\subseteq}^{(2)} \mathsf{P}(S) \overbrace{\subseteq}^{(1)} G \setminus S \tag{2.3}$$

Intuitively, this means (1) there are no edges between elements of S and (2) all elements attacking S are in turn attacked by S. We collect the set of all semikernels of G in the set $SK(\mathsf{G})$.

[2] $\mathcal{L}_\kappa$ denotes propositional language with formulae of finite depth, formed over an arbitrary set of atoms by unary negation and (possibly) infinite conjunctions of sets of formulae with cardinality $< \kappa$. Binary connectives, such as $\leftrightarrow$, are encoded (but could be added).

[3] For the theory of formal argumentation generally, see, e.g., [10, 15].

Theories and graphs can be transformed into each other, along with the associated models and kernels. A theory Γ in GNF gives rise to a graph $\mathcal{G}(\Gamma)$ with all atoms as vertices and edges from every x on the left-hand side of a GNF formula in Γ, to each y_i on its righ-hand side, i.e., $\mathsf{N} = \{\langle x, y_i \rangle \mid x \in G, i \in I_x\}$. E.g., Θ_1 from Example 2.2 has the graph $\mathcal{G}(\Theta_1): \circlearrowleft a \to b \rightleftarrows c$.

The theory of a graph $\mathsf{G} = \langle G, \mathsf{N} \rangle$ is

$$\mathcal{T}(\mathsf{G}) = \{x \leftrightarrow \bigwedge_{y \in \mathsf{N}(x)} \neg y \mid x \in G\}.$$

(When x is a *sink*, $\mathsf{N}(x) = \varnothing$, this becomes $x \leftrightarrow \top$, i.e., x is included in $\mathcal{T}(\mathsf{G})$.) The two are inverses, so we ignore usually the distinction between theories (in GNF) and graphs, viewing them as alternative presentations. Typically, Γ denotes such a theory or a graph, while G the corresponding set of atoms/vertices.

The presentations are equivalent also semantically: for corresponding graph and theory, the kernels of the former and models of the latter are in bijection. A kernel of a graph G can be defined equivalently as a partition α of G into two disjoint subsets $\langle \alpha^{\mathbf{1}}, \alpha^{\mathbf{0}} \rangle$ such that $\forall x \in G$:

$$\begin{array}{llll} (a) & x \in \alpha^{\mathbf{1}} & \Leftrightarrow & \forall y \in \mathsf{N}(x) : y \in \alpha^{\mathbf{0}} \\ (b) & x \in \alpha^{\mathbf{0}} & \Leftrightarrow & \exists y \in \mathsf{N}(x) : y \in \alpha^{\mathbf{1}}. \end{array} \tag{2.4}$$

Conditions (a) and (b) are equivalent for total α (with $\alpha^{\mathbf{0}} = G \setminus \alpha^{\mathbf{1}}$), so one will suffice, until we consider partial structures. A total α satisfies (2.4) iff $\alpha^{\mathbf{1}} \in Ker(\mathsf{G})$. On the other hand, satisfaction of (2.4) at every $x \in G$ is equivalent to the satisfaction of the respective GNF theory $\mathcal{T}(\mathsf{G})$. So, for corresponding graph and theory, we identify also kernels of the former and models of the latter.

Similarly, the notion of a semikernel can be defined equivalently as a partition α of G into three disjoint subsets $(\alpha^{\mathbf{1}}, \alpha^{\mathbf{0}}, \alpha^{\perp})$ such that $\forall x \in G$:

$$\begin{array}{llll} (a) & x \in \alpha^{\mathbf{1}} & \Rightarrow & \forall y \in \mathsf{N}(x) : y \in \alpha^{\mathbf{0}} \\ (b) & x \in \alpha^{\mathbf{0}} & \Leftrightarrow & \exists y \in \mathsf{N}(x) : y \in \alpha^{\mathbf{1}}. \end{array} \tag{2.5}$$

A threeway partition α satisfies (a) and (b) above iff $\alpha^{\mathbf{1}}$ is a semikernel. $\alpha^{\mathbf{1}}$ is a kernel iff α is a semikernel with $\alpha^{\perp} = \emptyset$. The reader interested in argumentation will notice the correspondence with labelling-based formulations of argumentation semantics [6, 8]. In argumentation theory, several additional semantics have been defined by placing constraints on semikernels (admissible sets). [5, 7, 10]. In the present paper, we will present a new semantics that adds a constraint on $\alpha^{\perp}$ (c.f., Eq. 6.1). Interestingly, the new semantics *uniquely* characterises the paradoxes (undecided elements) of G, while at the same time agreeing with the classical semantics on all graphs that admit stable sets. To our knowledge, this combination of desirable features is not found in any of the existing semantics for formal argumentation based on admissible sets. For further details, we refer the reader to Section 6 below, which is largely self-contained.

Example 2.6. *The graphs for the discourses from Example 2.2 are:*

$$\mathcal{G}(\Theta_1): \circlearrowleft a \to b \rightleftarrows c \qquad \mathcal{G}(\Theta_2): \circlearrowleft a \to b \rightleftarrows c \leftarrow d \circlearrowright$$

In $\mathcal{G}(\Theta_1)$, the partition $\alpha = \langle \{b\}, \{a, c\} \rangle$ is the only one satisfying (2.4), i.e., $\alpha^{\mathbf{1}} = \{b\}$ determines the only model of Θ_1/kernel of $\mathcal{G}(\Theta_1)$.

In $\mathcal{G}(\Theta_2)$, the same α satisfies (2.4) at $\{a, b, c\}$, but leaves no satisfying assignment at d. Letting, on the other hand, $\beta^{\mathbf{1}} = \{c\}$ and

$\beta^0 = \{b, d\}$ satisfies (2.4) at $\{b, c, d\}$, but leaves no possible assignment to a. michα^1 and β^1 are semikernels but the graph has no kernel, i.e., the discourse is paradoxical.

The two equivalent representations give two sets of means for analyzing each particular case. General patterns of (non-)paradoxes are naturally expressible as properties of graphs, which becomes increasingly popular [2, 9, 12–14]. Section 3 provides some examples and new statements. But when a particular case does not fall under any general pattern, actual reasoning can be necessary for adequate analysis. Sections 4-5 present a complete reasoning system, capable of demonstrating paradoxicality of arbitrary countable discourse, and describe some of its features.

3 GRAPH PATTERNS OF PARADOX

The observation that a language with negation and self-reference allows to formulate paradoxes, finds a graph counterpart: graphs containing odd cycles may fail to possess kernels. The simplest example is a single loop, corresponding to the liar. The following theorem of Richardson from 1953 formulates precisely the intuition of the opposite, namely, that only self-negating circularity is vicious. (A graph is *locally finite* if $\forall x \in G : \mathsf{N}(x)$ is finite. A *ray* is an inifinite, simple, outgoing path, i.e., $\{x_i \in G \mid i \in \omega\}$, where $\forall i, j \in \omega : i \neq j \Rightarrow x_i \neq x_j \wedge x_{i+1} \in \mathsf{N}(x_i)$. $\vec{\mathsf{G}}$ denotes all rays of G.)

Theorem 3.1 ([16]). *A graph without odd cycles has a kernel if it (i) is locally finite or (ii) has no rays.*

In particular, a finite discourse without odd cycles is never paradoxical. Thus the liar, the minimal odd cycle, is the paradigmatic finitary paradox. Kernel theory provides more detailed conditions for the exitence of kernels, admitting also odd cycles, see e.g., [4, 12].

The additional condition (i) or (ii) excludes infinitary paradoxes, for instance, the Yablo graph $\mathsf{Y} = \langle \mathbb{N}, < \rangle$, which appears equally paradigmatic for infintary paradoxes, as liar is for the finitary ones. One problem is to capture a general feature, responsible for making a "Yablo-like" graph paradoxical. The other is to prove that it actually is the source of paradox, like an odd cycle is in the finitary case.

Concerning the first issue, excluding Y as a topological minor does not suffice, as shown in [14]. A more general notion of digraph minor is presented in [2] without, however, any results showing that exclusion of such a Y-minor might guarantee the existence of a kernel. The formulation below appears thus to be the first accompanied by such results. (The rest of this section, including proofs, is elaborated in unpublished [17].)

A vertex r on a ray R *dominates* R if there are infinitely many mutually disjoint paths from r to R. For instance, in Y, every vertex dominates every ray on which it lies. Call a graph *safe* if it has no odd cycle and no ray with infinitely many vertices dominating it. (Exclusion of such rays is equivalent to the exclusion of Y-minor from [2].)

Conjecture 3.2. *A safe graph has a kernel.*

Proving this appears unexpectedly difficult, but important special cases do hold. Their formulation uses the notion of a graph's *end* as the subgraph reaching a specific ray, i.e., $\mathsf{P}^*(R)$ for any $R \in \vec{\mathsf{G}}$. Ends capture, so to speak, the ways in which a graph "proceeds towards infnity". In a graph with only one end, i.e., $\forall R, Q \in \vec{\mathsf{G}} : \mathsf{P}^*(R) = \mathsf{P}^*(Q)$, all rays are indefinitely reachable from each other: no matter how far we proceed along one, we can always switch to another. This is the case with Y which, having uncountably many tail-disjoint rays, has only one end. By contrast, consider a graph where from each tail of a ray R, there exits another ray which neither returns to R nor meets any other ray. The graph has infinitely many ends, as one can "proceed towards infinity" along infinitely many "different" ways, where two ways are "different" if one of them becomes eventually unreachable from the other. To prove the conjecture, what is needed is a corresponding generalisation of the following result.

Theorem 3.3 ([17]). *A safe graph with finitely many ends has a kernel.*

The theorem follows from more general statements, showing kernel existence for various safe graphs with infinitely many ends. For instance, safe graphs with countably many ends, where no end is a subset of another, have kernels. In particular, graphs with only countably many tail-disjoint rays, where no end is a subset of another, have kernels.

We also have the following generalization of point (i) from Theorem 3.1, admitting arbitrary number of ends. A vertex $x \in G$ is *finitely separable from (tails of) all rays* if there is a finite set $F \not\ni x$ such that there is no ray starting from x in the induced subgraph $G \setminus F$.[4] Finite separability generalizes local finiteness, admitting infinite branching as long as it does not lead to infinitely many disjoint rays. The condition below implies safety.

Theorem 3.4 ([17]). *If* G *has no odd cycle and every vertex is finitely separable from all rays, then* G *has a kernel.*

In view of these results, in particular, Theorem 3.3, one can claim Conjecture 3.2 to hold for most imaginable discourses. Problems with its unrestricted confirmation concern peculiar ways in which discourses can stretch "towards infinity" along infinitely many independent ways. These referential patterns are hard to formulate and analyse in purely logical terms. Graphs, providing means for this, leave the hope that also the general case can be proven.

4 INFINITARY RESOLUTION

Graph patterns give means for excluding many common cases of paradox and enhance the general understanding of its structure. But since the problem of kernel existence is difficult (NP-complete for finite graphs, Σ^1_1-complete for recursive ones and, in general, equivalent to consistencey of theories in infinitary propositional logic, [3]), one can hardly expect any simple sufficient criteria, which would be also necessary. Many graphs with odd cycles have kernels, as do many unsafe ones without odd cycles. It is therefore desirable to have also more detailed means for analyzing any particular case.

The inference system presented below provides such means. It is essentially (negative and positive) hyper-resolution, handling infinitary clausal theories arising from GNF. The two implications in (2.1) give two kinds of clauses for every $x \in G$:

OR-clause: $x \vee \bigvee_{i \in I_x} y_i$, written as $xy_1y_2...$

NAND-clauses: $\neg x \vee \neg y_i$, for every $i \in I_x$, denoted $\overline{xy_i}$.

[4] For $H \subseteq G$, the subgraph of $\langle G, \mathsf{N} \rangle$ induced by H is $\langle H, \mathsf{N} \cap (H \times H) \rangle$.

In terms of a graph, its theory contains, for every $x \in G$, the OR-clause $\mathsf{N}[x] = \{x\} \cup \mathsf{N}(x)$ and for every $y \in \mathsf{N}(x)$, the NAND-clause $\overline{xy}$. For the graphs from Example 2.6, the clausal theories are:

$\Theta_1' = \{ab, bc, \overline{ab}, \overline{bc}, \overline{a}\}$ and $\Theta_2' = \{ab, bc, cd, \overline{ab}, \overline{bc}, \overline{cd}, \overline{a}, \overline{d}\}$.
We treat both kinds of clauses as sets of atoms, and overbars mark only that a set is a NAND-clause. We can therefore write, e.g., $\overline{xy} \subseteq G$. A set $A \subseteq G$ is also an OR-clause, $\overline{A} = \{\overline{a} \mid a \in A\}$ a NAND-clause. Sets of unary clauses are denoted $A^+ = \{\{a\} \mid a \in A\}$ and $A^- = \{\{\overline{a}\} \mid a \in A\}$. The considered language contains only OR and NAND clauses, but no mixed ones.

Of primary interest to us are graphs (GNF theories) but several results hold for theories with finite NAND-clauses. Saying "every Γ", we mean such theories. The following system RIP is complete for such theories with countable OR set, denoted C-F, while it is sound for arbitrary theories (also with infinite NANDs, which we do not consider.) Proofs missing below can be found in [18].

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

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

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

Proofs are well-founded trees with (Ax) at the leafs, rule applications at all internal nodes, and the conclusion at the root. In particular, every branch of a proof is finite. There are no cardinality restrictions on the index sets I and finitary logic is a special case.

The rule (Rneg) derives a NAND from NANDs, using a single OR as a side formula, while (Rpos) derives an OR from ORs, using NANDs as side formulae. In (Rneg), $\overline{a_i A_i}$ denotes the NAND $\overline{\{a_i\} \cup A_i}$, where A_i may be empty. These negative premises are "joined" – into the union of all $\overline{A_i}$ – by the OR-clause O, with each $a_i \in O$ belonging to one $\overline{a_i A_i}$.

In (Rpos), among the OR-premises there is the "main" clause A, containing a subset $\{a_i \mid i \in I\}$ such that for each a_i, there is an OR-premise $B_i K_i$ ($B_i \cup K_i$), with side premises $\overline{a_i k}$ for all $k \in K_i$. The conclusion joins the OR-clauses removing the atoms from the negative premises. A special case of the rule has only the main OR-premise A with the side premises $\Gamma \vdash \overline{a}_i, i \in I$, yielding the conclusion $A \setminus \{a_i \mid i \in I\}$.

As a simple example of diagnosing the paradox by proving the empty clause $\{\}$, consider the Yablo graph $\mathsf{Y} = \langle \mathbb{N}, < \rangle$, where ORs are $O_i = \{j \mid j \geq i\}$ for all $i \in \mathbb{N}$, and NANDs all pairs $\overline{ij}$, for $i \neq j$. For each i, starting with the axioms $\overline{ij}$ for all $j > i$ and using O_{i+1}, yields $\overline{i}$, and from these $\{\}$ follows using O_1:

$$\frac{\dfrac{\overline{12}, \overline{13}, \overline{14}, \ldots}{\overline{1}} \; O_2 \quad \dfrac{\overline{23}, \overline{24}, \overline{25}, \ldots}{\overline{2}} \; O_3 \;\; \ldots \;\; \dfrac{\{\overline{ij} \mid j > i\}}{\overline{i}} \; O_{i+1} \;\; \ldots}{\{\}} \; O_1$$

RIP contains two independent systems:

(Neg) consisting of (Ax) and (Rneg), and

(Pos) consisting of (Ax) and (Rpos).

Each system is sound (for all theories) and refutationally complete (for C-F-theories) on its own.

THEOREM 4.1 ([18]). *For every $C \subseteq G$ and*
1. for every $\Gamma : (\Gamma \vdash_{Neg} \overline{C} \vee \Gamma \vdash_{Pos} \overline{C}) \Rightarrow \Gamma \models \overline{C}$,
2. for C-F $\Gamma : Mod(\Gamma) = \varnothing \Rightarrow (\Gamma \vdash_{Neg} \{\} \wedge \Gamma \vdash_{Pos} \{\})$,
3. for C-F $\Gamma : \Gamma \models \overline{C} \Leftrightarrow \Gamma \cup C^+ \vdash_{Neg} \{\}$ and $\Gamma \models C \Leftrightarrow \Gamma \cup C^- \vdash_{Pos} \{\}$.

1 gives that an arbitrary discourse Γ is paradoxical if $\Gamma \vdash \{\}$, while 2 that if a paradoxical Γ is C-F, then $\Gamma \vdash \{\}$. We note also the following corollary, reflecting well the intuition of paradox as a statement appearing (for reasoning) both true and false:

$$\Gamma \vdash \{\} \Leftrightarrow \exists x : \Gamma \vdash x \wedge \Gamma \vdash \overline{x} \qquad (\textit{denoted } \Gamma \vdash \bot(x)).$$

5 NONEXPLOSIVENESS

The rules are essentially classical hyper-resolution and RIP is refutationally complete for classical semantics (of C-F theories, Theorem 4.1.3). For diagnosing paradox of a discourse Γ, however, the refutational reasoning is not used, since one simply tries to derive $\{\}$ from Γ. If we now focus on such a direct (nonrefutational) derivability, i.e., ask if $\Gamma \vdash C$, instead of $\Gamma, C^- \vdash \{\}$, there emerge unexpected, paraconsistent and relevant features of the system.

First, completeness becomes restricted to some nonredundant clauses. ($\widetilde{\ }$ denotes, in a given context, either everywhere positive or everywhere negative occurrences.)

COROLLARY 5.1 ([18]). *For C-F Γ and $A \subseteq G : \Gamma \models \widetilde{A} \Leftrightarrow \exists B \subseteq A : \Gamma \vdash \widetilde{B}$.*

Consequently, the logic does not have weakening – hence neither *Ex Falso Quodlibet*. Its nonexplosiveness gives a paraconsistent ability to contain paradox and reason – classically – about the subdiscourse unaffected by it.

As a simplest example, for $\Gamma = \{x, \overline{x}, y\}$, we have $\Gamma \vdash \{\}$ but also $\Gamma \nvdash \overline{y}$. Looking at its graph – $x \circlearrowleft \quad y$ – this is as it should be: paradox at x is in no way "connected" to y. This is the essence of the phenomenon, which we now describe in more detail.

EXAMPLE 5.2. *The closure of $y \to z \to x \circlearrowleft$ contains, besides $\{\}$, all literals. Provability of both x and $\overline{x}$, i.e., the paradox at x, pollutes the whole discourse.*

The discourse $\{yz, \overline{yz}, zxs, \overline{zx}, \overline{zs}, x, \overline{x}, s\}$, i.e., $s \leftarrow y \to z \to x \circlearrowleft$, is still paradoxical at x and $\{\}$ is still provable, but neither is $\overline{y}$ nor z. The closure contains only the literals $\{x, \overline{x}, s, \overline{z}, y\}$, showing that x is the only problem, which does not affect the rest of the discourse.

To identify semantic counterpart of this nonexplosiveness, we introduce the operation which, for $\Gamma \subseteq \mathcal{P}(Y)$ and $X \subseteq Y$, removes all atoms X from all clauses of Γ, removing also the empty clause, if it appears in the process (a negative clause $\overline{A}$ is just negatively marked set A – the operation removes atoms from both positive and negative clauses):

$$\Gamma \setminus\!\!\setminus X = \{C \setminus X \mid C \in \Gamma\} \setminus \{\{\}\}.$$

Let us now denote:

$G^{\bot} = \{x \in G \mid \Gamma \vdash x \wedge \Gamma \vdash \overline{x}\}$
$\Gamma^{ok} = \Gamma \setminus\!\!\setminus G^{\bot} = \{C \setminus G^{\bot} \mid C \in \Gamma\} \setminus \{\{\}\}$
$G^{ok} = G \setminus G^{\bot} = \bigcup \Gamma^{ok}$.

$G^{\bot}$ contains all statements involved in the paradox and the story ends here when it is empty or covers the whole G. But otherwise Γ^{ok} remains consistent alongside $G^{\bot}$ and conservative over Γ with respect to nonparadoxical facts $x \in G^{ok}$, in the following sense.

FACT 5.3 ([18]). *For C-F Γ with $G^{ok} \neq \varnothing$:*

(1) $\forall \widetilde{D} \in \Gamma^{ok} : \Gamma \vdash \widetilde{D}$, *so* $\forall C \subseteq G^{ok} : \Gamma^{ok} \vdash \widetilde{C} \Rightarrow \Gamma \vdash \widetilde{C}$.
(2) $\Gamma^{ok} \nvdash \{\}$.
(3) $\forall x \in G^{ok} : \Gamma^{ok} \vdash x \Leftrightarrow \Gamma \vdash x$ *and* $\Gamma^{ok} \vdash \overline{x} \Leftrightarrow \Gamma \vdash \overline{x}$.
(4) $\exists x \in G^{ok} : \Gamma^{ok} \nvdash \overline{x}$, *hence also* $\Gamma \nvdash \overline{x}$.
(5) $\forall x \in G^{ok} : \Gamma^{ok} \nvdash \overline{x} \Rightarrow \mathsf{N}(x) \cap G^{\perp} = \varnothing$ *(when* Γ *is a graph).*

For a graph Γ, Γ^{ok} is almost the theory of its subgraph induced by G^{ok}, except for some differences at its border vertices $brd(G^{ok}) = \{x \in G^{ok} \mid \mathsf{N}(x) \nsubseteq G^{ok}\}$.

Example 5.4. *Consider the following* Γ : $y \twoheadrightarrow z \twoheadrightarrow x \quad s \rightleftarrows t$.

$$\begin{array}{lcl}
\Gamma & = & \{yz, \overline{yz}, zxs, \overline{zx}, \overline{zs}, st, \overline{st}, x, \overline{x}\} \\
G^{\perp} & = & \{x\} \\
G^{ok} & = & \{y, z, s, t\} \\
\Gamma^{ok} & = & \{yz, \overline{yz}, zs, \overline{z}, \overline{zs}, st, \overline{st}\} \\
brd(G^{ok}) & = & \{z\}
\end{array}$$

The subgraph induced by G^{ok} *is* $\underline{\Gamma}^{ok} : y \twoheadrightarrow z \twoheadrightarrow s \rightleftarrows t$ *and its theory*

$$\mathcal{T}(\underline{\Gamma}^{ok}) \quad = \quad \{yz, \overline{yz}, zs, \overline{zs}, st, \overline{st}\},$$

while Γ^{ok} *contains, in addition,* $\overline{z}$.

Border vertices enter as such negative clauses into

$$\Gamma^{ok} = \mathcal{T}(\underline{\Gamma}^{ok}) \cup (brd(G^{ok}))^{-},$$

so we can view Γ^{ok} as the subgraph $\underline{\Gamma}^{ok}$ induced by G^{ok}, with a new loop at each border vertex. It is not paradoxical, Fact 5.3.2, and its models are kernels of $\underline{\Gamma}^{ok}$ excluding border vertices:

$$Mod(\Gamma^{ok}) = \{L \in Ker(\underline{\Gamma}^{ok}) \mid brd(G^{ok}) \subseteq \mathsf{P}(L)\}.^{5}$$

In the above example, $\underline{\Gamma}^{ok}$ has two kernels $\{t, z\}$ and $\{s, y\}$, but only the latter gives a model of Γ^{ok}, which requires $\overline{z}$, i.e., exclusion of z.

This semantics does not relate RIP to G, but to G^{ok}, obtained from G by removing elements that are provably paradox. Hence, the semantics is defined in terms of the reasoning system, a dependence we would like to get rid of by relating provability directly to an independently formulated structure found in G.

To obtain this, we show that the relation between the paradoxical graph Γ and Γ^{ok} can be specified further in semantic terms, based on the notion of a semikernel. This will also yield a new semantics for argumentation frameworks based on admissible sets (corresponding to semikernels).

6 A NEW SEMANTICS FOR PROPOSITIONAL DISCOURSES AND ARGUMENTATION FRAMEWORKS

Unlike previous semantics for discourses and argumentation frameworks, the system RIP encodes the following intuition about paradoxical statements, formulated here in terms of argumentation: if A attacks B and B is paradoxical, then A is also paradoxical. In other words, if you attack a paradox and fail, you become part of the paradox. This follows naturally from the intuition that a paradox is something we must try to avoid: if B, referring to A, is paradoxical, then A could possibly resolve it; but since it does not, it too is involved in this paradox.[6] If reasoning proceeds on this premise, it is only natural that a failed attack on a paradox will cause the paradox to spread to its attacker. As we will prove in the following, this semantic idea is the key to characteristing RIP directly in terms of G.

To make this precise, we first adopt the labelling-based approach, as it closely tracks the intuition presented above. Specifically, it follows from what we have said that B should be labelled as a paradox if, and only if, everything that attacks B is also a paradox. Formally, α is a *PS* labelling of G if it partitions G into three disjoint subsets $\langle \alpha^{\mathbf{1}}, \alpha^{\mathbf{0}}, \alpha^{\perp} \rangle$ such that $\forall x \in G$:

$$\begin{array}{llcl}
(a) & x \in \alpha^{\mathbf{1}} & \Rightarrow & \forall y \in \mathsf{N}(x) : y \in \alpha^{\mathbf{0}} \\
(b) & x \in \alpha^{\perp} & \Leftrightarrow & \forall y \in \mathsf{N}(x) : y \in \alpha^{\perp} \\
(c) & x \in \alpha^{\mathbf{0}} & \Leftrightarrow & \exists y \in \mathsf{N}(x) : y \in \alpha^{\mathbf{1}}
\end{array} \qquad (6.1)$$

In view of these conditions, it follows that all PS labellings are semikernel labellings, c.f. Eq. 2.5. The opposite does not hold in general, as illustrated by the following graph: $f \twoheadrightarrow y \twoheadrightarrow s$. The only PS labelling of this graph is α with $\alpha^{\mathbf{1}} = \alpha^{\mathbf{0}} = \emptyset$ and $\alpha^{\perp} = \{f, y, s\}$. This illustrates that PS labellings can designate *unattacked* elements as paradoxes. This might be surprising, but follows from the fact that paradoxes spread to unsuccessful attackers. Specifically, even if s is not attacked, its attack on y cannot succeed. This is because y attacks the paradox f, which necessarily spreads to y since y cannot be succesful. Hence, regarding y as having been defeated by s is not possible and the paradox spreads further to s. Recognise that if the "base" paradox f could be defeated, this behaviour would not occur; after all, any kernel α is trivially also a PS labelling.

The underlying reason why PS labellings behave in the way described above is that they satisfy a closure property. Specifically, for all α, if α is a PS labelling then for all $x \in G$:

$$\begin{array}{ll}
(i) & x \in \alpha^{\perp} \Rightarrow \mathsf{N}^{*}(x) \subseteq \alpha^{\perp} \\
(ii) & x \in \alpha^{\mathbf{1}} \cup \alpha^{\mathbf{0}} \Rightarrow \mathsf{P}(x) \subseteq \alpha^{\mathbf{1}} \cup \alpha^{\mathbf{0}}
\end{array} \qquad (6.2)$$

Point (i) above follows from condition (b) of Eq. 6.1 while point (ii) is equivalent to point (i) (recall that α is a threeway partition). If α is a semikernel labelling, then point (ii) also implies condition (b) of Eq. 6.1, as the reader can verify. Hence, we arrive at a characterisation of PS labellings in set-theoretic terms. Let $\mathsf{P}[S] = S \cup \mathsf{P}(S)$. Then we say that S is a PS set if S is a semikernel and $\mathsf{P}(\mathsf{P}[S]) \subseteq \mathsf{P}[S]$. That is, a PS set is a P-closed semikernel. It follows that S is a PS set if, and only if, α_S is a PS labelling, where $\alpha_S^{\mathbf{1}} = S, \alpha_S^{\mathbf{0}} = \mathsf{P}(S)$ and $\alpha_S^{\perp} = G \setminus (\alpha_S^{\mathbf{1}} \cup \alpha_S^{\mathbf{0}})$. We collect the PS sets of G in the set $PS(\mathrm{G})$.

To characterise models of G^{ok} in terms of G, we need to restrict attention to maximal PS sets. Specifically, α is an *mPS* labelling if α is a PS labelling that also satisfies the following constraint:

$$\forall \beta \in PS(\mathrm{G}) : \beta^{\mathbf{0}} \cup \beta^{\mathbf{1}} \subseteq \alpha^{\mathbf{0}} \cup \alpha^{\mathbf{1}} \qquad (6.3)$$

Similarly, S is said to be an mPS set if $\alpha_S = \langle S, \mathsf{P}(S), G \setminus \mathsf{P}[S] \rangle$ is an mPS labelling. $mPS(\mathrm{G})$ denotes all mPS sets of G.

We now demonstrate a basic property of mPS semantics, namely that every mPS labelling has the same paradoxical domain. This is satisfying for both technical and conceptual reasons. Specifically, it

[5] This makes sense as $\forall b \in brd(\underline{\Gamma}^{ok}) : b \notin sinks(\underline{\Gamma}^{ok})$, since $\varnothing \neq \mathsf{N}(b) \subseteq G^{\perp} \overset{6.7}{\Rightarrow} b \in G^{\perp}$.

[6] While paradoxes are often referred to as undecided arguments in argumentation theory, there is still a clear sense that the cause of indeterminacy is that certain arguments are paradoxical due to vicious circularity, a form of argument interaction that should be avoided [1].

shows that mPS semantics is capable of uniquely diagnosing paradoxes, while at the same time agreeing with the classical semantics (kernels/stable sets) on all graphs that admit kernels. To our knowledge, the mPS semantics is the first semantics for propositional discourses and argumentation frameworks that has this property.[7] The proof of the following lemma can be found in the appendix.

Lemma 6.4. *For all* G *and all mPS labellings* α *and* β, $\alpha^{\perp} = \beta^{\perp}$.

It follows that maximising $\alpha^{\mathbf{1}}$ is the same as minimising $\alpha^{\perp}$, for all PS labellings α. For admissible labellings, this is not the case, giving rise to the distinction between preferred semantics (maximising $\alpha^{\mathbf{1}}$) [10] and semi-stable semantics (minimising $\alpha^{\perp}$) [5, 7]. In our view, the fact that no corresponding distinction arises when we start from PS labellings is a strength; maximising acceptability *should* coincide with minimising indeterminacy, at least as long as avoiding paradox is regarded as a principle of acceptable reasoning.

Lemma 6.4 also shows that mPS sets are a faithful generalisation of kernels and stable sets, arguably more so than any semantics previously considered. Every semikernel is trivially a kernel in the subgraph induced by the semikernel and its in-neigbourhood, but the mPS semantics satisfies a much stronger property: all mPS sets are kernels of the *same* induced subgraph of G. Specifically, let G^- denote the subgraph induced by $G \setminus \alpha^{\perp}$ for some arbitrary mPS labelling α. Then $\beta^{\mathbf{1}}$ is a kernel for G^- for each mPS labelling β. Showing that $\mathsf{G}^- = \mathsf{G}^{ok}$ is then the key to our main result, characterising the semantics for RIP directly in terms of G. Basically, we have to show that the provable paradoxes correspond to the paradoxes discoverd by the mPS semantics. The proof of this involves some auxiliary structures and can be found in the appendix.

Theorem 6.5. *For countable* G :

$$Mod(\Gamma^{ok}) = \{K \in sol(\mathsf{G}^{ok}) \mid brd(\mathsf{G}^{ok}) \subseteq \mathsf{P}(K)\} = mPS(\mathsf{G}).$$

The semantics $Mod(\Gamma^{ok})$ – addressing only a part of Γ – explains the nonexplosive behavior of RIP: reasoning from Γ is sound also for partial structures. Besides contrarieties $\perp(x)$, provable when $G^{\perp} \neq \varnothing$, RIP proves neither simply facts true in all kernels of Γ (as does classical logic), nor simply facts implied by all its semikernels (as does **Ł3**, [12]), but facts true in maximal P-closed semikernels, namely mPS sets. For literals (in countable graphs), this is Fact 5.3.3, while the following implies the general case for arbitrary graphs (inclusion to the left holds also for Γ with infinite NANDs, but to the right requires finite NANDs).

Theorem 6.6 ([18]). *For every* Γ : $Mod(\Gamma^{ok}) = Mod(Th(\Gamma)|_{G^{ok}})$, *where* $Th(\Gamma)|_{G^{ok}} = \{\widetilde{C} \subseteq G^{ok} \mid \Gamma \vdash \widetilde{C}\}$.

6.1 Propagation of paradox

As we have seen, paradox spreads to all its attackers. This may seem surprising, since reading a path from x to y as x "referring to" or "depending on" y, a paradox pollutes thus everything on which it depends. For instance, in "This statement is false and the sun is a star", i.e., $f \to y \to s$, the paradox at f "refers to" the sink s (claiming that the sun is a star) and s becomes a paradox for preventing y (denying that the sun is a star) from resolving f

This is not to suggest that "the sun is a star" is paradoxical on its own, only that it gives the paradoxical whole when combined with the contingent liar paradox (f). Like consistency, paradox is genuinely holistic. Or to put it differently, the claim that the sun is a star *becomes* paradoxical in the present discourse, because it is used to reinstate a liar paradox, defending such a paradox against one of its attackers. To "repair" this anomaly, removing the loop at f is as good as removing s.

Paradox can also spread upwards, along P, as in $x \leftarrow z$, where provability of $\perp(x)$ leads to provability of $\perp(z)$. But such upward propagation can be interrupted. In Example 5.4, $G^{ok} = \{y, z, s, t\}$ – both z and y "depend" on the paradox at x, but are not affected by it.

A sufficient condition for an upward propagation of paradox is that all paths from a given statement reach, eventually, a paradox. A *complete path* is a path (i.e., $\pi \in G^I$ with $I \in \omega^+$ and $\pi_{i+1} \in \mathsf{N}(\pi_i)$ for all $i+1 \in I$) which is infinite or terminates with a sink. $paths(x)$ denotes all paths starting from x.

Fact 6.7 ([18]). *For an* x *in any graph* Γ, *if every complete* $\pi \in paths(x)$ *contains a paradoxical* π_i, *i.e.,* $\Gamma \vdash \perp(\pi_i)$, *then* $\Gamma \vdash \perp(x)$.

Γ from Example 5.4 illustrates thus the only possibility of preventing the propagation of paradox upwards by some path which, exiting from a border vertex, like $z \in brd(\Gamma^{ok})$, meets no paradox and forces z to be false.

7 CONCLUDING REMARK

Like in logics with internal T-predicate, paradox formulated in GNF becomes a special case of inconsistency: a discourse is paradoxical when the T-schemata of its statements are inconsistent. The graph representation gives a precise grasp of vicious circularities, confiriming, for instance, the intuition that negative self-reference is necessary (and not only sufficient) for finitary paradoxes. Furthermore, a general condition, conjectured necessary for occurrence of every paradox, is verified for a wide class of graphs.

The reasoning system RIP allows to analyse particular cases which do not fall under any general pattern of (non)paradox. The analysis can, besides diagnosing paradox, identify the nonparadoxical subdiscourse and its classical concequences, which are not affected by the surrounding inconsistency. Statements involved into paradox become thus identifiable as being both provably true and provably false. This paraconsistent effect arises from nonrefutational use of hyper-resolution, which deviates from classical reasoning only by the exclusion of weakening. Semantically, such statements fall out of maximal P-closed semikernels, providing the semantics which coincides with the classical one for the nonparadoxical discourses and, in general, amounts to the classical interpretation of the maximal nonparadoxical subdiscourse.

REFERENCES

[1] T. J. M. Bench-Capon. 2016. Dilemmas and paradoxes: cycles in argumentation frameworks. *J Logic Computation* 26 (2016), 1055–1064. Issue 4.

[2] Timo Beringer and Thomas Schindler. 2017. Graph-theoretic analysis of semantic paradoxes. *[submitted]* (2017).

[3] Marc Bezem, Clemens Grabmayer, and Michal Walicki. 2012. Expressive power of digraph solvability. *Annals of Pure and Applied Logic* 163, 3 (2012), 200–212.

[7] Another argumentation semantics that unqiuely characterises paradoxes is the so-called *grounded semantics* [10], which can be defined as the minimal complete set (it is easily shown to be unique). Unlike the mPS semantics, however, the grounded semantics does not agree with the classical semantics on non-paradoxical frameworks.

[4] Endre Boros and Vladimir Gurvich. 2006. Perfect graphs, kernels and cooperative games. *Discrete Mathematics* 306 (2006), 2336–2354.

[5] Martin W.A. Caminada. 2006. Semi-Stable Semantics. In *Proceedings of the 2006 conference on Computational Models of Argument: Proceedings of COMMA 2006.* IOS Press, Amsterdam, The Netherlands, The Netherlands, 121–130.

[6] Martin W. A. Caminada. 2006. On the Issue of Reinstatement in Argumentation. In *Logics in Artificial Intelligence, 10th European Conference, JELIA 2006, Liverpool, UK, September 13-15, 2006, Proceedings (Lecture Notes in Computer Science)*, Michael Fisher, Wiebe van der Hoek, Boris Konev, and Alexei Lisitsa (Eds.), Vol. 4160. Springer, 111–123.

[7] Martin W. A. Caminada, Walter Alexandre Carnielli, and Paul E. Dunne. 2012. Semi-stable semantics. *J. Log. Comput.* 22, 5 (2012), 1207–1254.

[8] Martin W. A. Caminada, Samy Sá, João Alcântara, and Wolfgang Dvorák. 2015. On the equivalence between logic programming semantics and argumentation semantics. *Int. J. Approx. Reasoning* 58 (2015), 87–111.

[9] Roy Cook. 2004. Patterns of paradox. *The Journal of Symbolic Logic* 69, 3 (2004), 767–774.

[10] Phan Minh Dung. 1995. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence* 77 (1995), 321–357.

[11] Sjur Dyrkolbotn. 2013. On a Formal Connection between Truth, Argumentation and Belief. In *Pristine Perspectives on Logic, Language, and Computation - ESSLLI 2012 and ESSLLI 2013 Student Sessions. Selected Papers (Lecture Notes in Computer Science)*, Margot Colinet, Sophia Katrenko, and Rasmus K. Rendsvig (Eds.), Vol. 8607. Springer, 69–90. DOI: http://dx.doi.org/10.1007/978-3-662-44116-9_6

[12] Sjur Dyrkolbotn and Michal Walicki. 2014. Propositional Discourse Logic. *Synthese* 191, 5 (2014), 863–899.

[13] Haim Gaifman. 1992. Pointers to truth. *The Journal of Philosophy* 89, 5 (1992), 223–261.

[14] Landon Rabern, Brian Rabern, and Matthew Macauley. 2013. Dangerous reference graphs and semantic paradoxes. *Journal of Philosophical Logic* 42 (2013), 727–765.

[15] Iyad Rahwan and Guillermo R. Simari. 2009. *Argumentation in Artificial Intelligence* (1st ed.). Springer.

[16] Moses Richardson. 1953. Solutions of irreflexive relations. *The Annals of Mathematics, Second Series* 58, 3 (1953), 573–590.

[17] Michal Walicki. 2017. Kernels of digraphs with finitely many ends. (2017). [available from http://www.ii.uib.no/~michal/].

[18] Michal Walicki. 2017. Resolving infinitary paradoxes. *Journal of Symbolic Logic* (2017). [to appear; available from http://www.ii.uib.no/~michal/].

APPENDIX

Proof of Lemma 6.4

For the proof of the main result, we use the following simple fact about semikernels.

FACT 7.1. *For all graphs* G*,*

(1) *If S is a semikernel and $T \subseteq S$, then $S \cap \mathsf{N}^*[T]$ is a semikernel, where $\mathsf{N}^*[T] = \mathsf{N}^*(T) \cup T$.*

(2) *If S and T are semikernels and $\mathsf{P}(S) \subseteq G \setminus T$, then $S \cup T$ is a semikernel.*

In view of this fact and Eq. 6.2.(ii), distinct PS sets, which disagree on at least one paradoxical element, can be combined as in the proof of the following lemma.

LEMMA 6.4. *For all* G *and all mPS labellings α and β, $\alpha^\perp = \beta^\perp$.*

PROOF. Let α and β be arbitrary mPS labellings in G with $\alpha^\perp \neq \beta^\perp$. We define $Q = \beta^{\mathbf{1}} \cap \alpha^\perp$ and $R = \beta^{\mathbf{0}} \cap \alpha^\perp$. We derive a contradiction from the following sequence of claims, which establish that $S = \alpha^{\mathbf{1}} \cup Q$ is a P-closed semikernel (with $Q \neq \emptyset$).

(a) $R \subseteq \mathsf{N}(Q)$, by $\beta^{\mathbf{1}}$ being a semikernel and Eq. 6.2.(i).
(b) $\mathsf{N}^*(Q) \subseteq \alpha^\perp$, by Eq. 6.2.(i).
(c) Q is a semikernel, by (b) and Fact 7.1.(1).
(d) $\mathsf{P}(Q) \subseteq G \setminus \alpha^{\mathbf{1}}$, by $Q \subseteq \alpha^\perp$ and $\alpha^{\mathbf{1}}$ being a semikernel (so that $\mathsf{P}(\alpha^\perp) \cap \alpha^{\mathbf{1}} = \emptyset$).
(e) S is a semikernel, by Fact 7.1.(2) (applicable by (c)-(d) above).
(f) S is P-closed, i.e., $\mathsf{P}(\mathsf{P}[S]) \subseteq \mathsf{P}[S]$. If $x \in S \subseteq \mathsf{P}[S]$, then trivially $\mathsf{P}(x) \subseteq \mathsf{P}[S]$. If $x \in \mathsf{P}(S)$, we have two cases.
 (i) $x \in \mathsf{P}(\alpha^{\mathbf{1}})$. Since α is a PS labelling: $\mathsf{P}(x) \cap \alpha^\perp = \emptyset$. Since $G \setminus \mathsf{P}[S] \subseteq \alpha^\perp : \mathsf{P}(x) \subseteq \mathsf{P}[S]$, as desired.
 (ii) $x \in \mathsf{P}(Q)$. Since α is a PS labelling and β is a semikernel labelling, we have $\mathsf{P}(Q) \subseteq \beta^{\mathbf{0}} \cap (\alpha^\perp \cup \alpha^{\mathbf{0}})$.
 If $x \in \alpha^{\mathbf{0}}$ then $\mathsf{P}(x) \subseteq \alpha^{\mathbf{1}} \cup \alpha^{\mathbf{0}}$, because α is a PS labelling. Hence $\mathsf{P}(x) \subseteq \mathsf{P}[S]$. If $x \in \beta^{\mathbf{0}} \cap \alpha^\perp$ then $\mathsf{P}(x) \subseteq (\beta^{\mathbf{1}} \cup \beta^{\mathbf{0}}) \cap (\alpha^\perp \cup \alpha^{\mathbf{0}})$. For any $y \in \mathsf{P}(x)$:
 $y \in (\beta^{\mathbf{1}} \cup \beta^{\mathbf{0}}) \cap \alpha^{\mathbf{0}} \subseteq \mathsf{P}(\alpha^{\mathbf{1}}) \subseteq \mathsf{P}(S)$, or
 $y \in \beta^{\mathbf{1}} \cap \alpha^\perp = Q \subseteq \mathsf{P}[S]$, or
 $y \in \beta^{\mathbf{0}} \cap \alpha^\perp$ – then $\exists z \in \beta^{\mathbf{1}} : y \in \mathsf{P}(z)$, since $\beta^{\mathbf{1}}$ is a semikernel. Since $z \in \mathsf{N}(y)$, so $z \in \alpha^\perp$ by Eq. 6.2.(i), which means that $z \in Q$ so that $y \in \mathsf{P}(Q) \subseteq \mathsf{P}(S)$.

Thus $S = \alpha^{\mathbf{1}} \cup Q$ is a P-closed semikernel. By point (a) and the fact that $\alpha^\perp \neq \beta^\perp$, it follows that $Q \neq \emptyset$, contradicting that α is an mPS labelling (Eq. 6.3). □

Proof of Theorem 6.5

To prove the characterisation theorem, we first define what it means for an mPS labelling α to satisfy a clause $\tilde{C}$. Ultimately, we want to use mPS labellings to characterise the two-valued (partial) semantics for RIP, so the generalisation of semantic entailment to three-valued labellings is used here only as an auxiliary structure. As noted in Section 6, the key to the proof of Theorem 6.5 is to show that the set of provable paradoxes coincides with the set of paradoxes identified under mPS semantics. In order to do so, it will be very useful to establish soundness of RIP with respect to the following notion of semantic entailment for three-valued labellings:

$$\begin{array}{lcl} \alpha \models C & \Leftrightarrow & C \cap \alpha^{\mathbf{1}} \neq \varnothing \vee C \subseteq \alpha^\perp \\ \alpha \models \overline{C} & \Leftrightarrow & C \cap \alpha^{\mathbf{0}} \neq \varnothing \vee C \subseteq \alpha^\perp \\ \alpha \models \{\} & \Leftrightarrow & \alpha^\perp \neq \varnothing \end{array} \tag{7.2}$$

CLAIM 7.3. RIP *is sound with respect to the three-valued semantic entailment as defined in Eq. (7.2).*

PROOF. Let $\alpha = \langle \alpha^{\mathbf{1}}, \alpha^{\mathbf{0}}, \alpha^\perp \rangle$ be an arbitrary structure (partition of G). If $\alpha \models \Gamma$ then obviously $\alpha \models C$ for each $C \in \Gamma$.

Let $\alpha = \langle \alpha^{\mathbf{1}}, \alpha^{\mathbf{0}}, \alpha^\perp \rangle$ be an arbitrary structure (partition of G). If $\alpha \models \Gamma$ then obviously $\alpha \models C$ for each $C \in \Gamma$.

(Rneg) If each $a_i A_i \subseteq \alpha^\perp$, then $\bigcup_i A_i \subseteq \alpha^\perp$. Otherwise, let $\varnothing \neq I_0 \subseteq I$ be such that $\forall i \in I_0 : a_i A_i \cap \alpha^{\mathbf{0}} \neq \varnothing$, while $\forall j \in I \setminus I_0 : a_j A_j \subseteq \alpha^{\mathbf{1}}$.

If $\forall i \in I_0 : a_i \notin \alpha^{\mathbf{1}}$ then $\{a_i \mid i \in I\} = \{a_i \mid i \in I_0\} \cup \{a_j \mid j \in J_0\}$ and neither of these two subsets intersects $\alpha^{\mathbf{1}}$, so that $\alpha \not\models \{a_i \mid i \in I\}$. Hence, $\exists i \in I_0 : a_i \in \alpha^{\mathbf{1}}$ and then $A_i \cap \alpha^{\mathbf{0}} \neq \varnothing$.

(Rpos) If $A \subseteq \alpha^\perp$ and all $B_i K_i \subseteq \alpha^\perp$, then the conclusion $C \subseteq \alpha^\perp$. Likewise, the conclusion $C \subseteq \alpha^\perp$ follows if $A \setminus \{a_i \mid i \in I\} \subseteq \alpha^\perp$ and all $B_i \subseteq \alpha^\perp$.

Otherwise, let us see if it is possible that $C \setminus \alpha^\perp \subseteq \alpha^{\mathbf{0}}$, when the premises are satisifed. We would then have that $\{a_i \mid i \in I\} \cap \alpha^{\mathbf{1}} \neq \varnothing$ and each $K_i \cap \alpha^{\mathbf{1}} \neq \varnothing$. So let $a_{i_0} \in \alpha^{\mathbf{1}}$ be a witness to the first. Then for all $k \in K_i$, we must have $k \in \alpha^{\mathbf{0}}$ to satisfy the third part of the premise. But then $B_i K_i \cap \alpha^{\mathbf{1}} = \varnothing$ and $B_i K_i \nsubseteq \alpha^\perp$. □

We are ready for the main proof, where we will use the notation $\mathrm{G}^{\mathbf{1}} = \{x \in G \mid \Gamma \vdash x\}$, $\mathrm{G}^{\mathbf{0}} = \{x \in G \mid \Gamma \vdash \overline{x}\}$, $\mathrm{G}^\perp = \mathrm{G}^{\mathbf{1}} \cap \mathrm{G}^{\mathbf{0}} = \{x \in G \mid \Gamma \vdash x \wedge \Gamma \vdash \overline{x}\}$, and for an $S \subseteq G : \alpha_S = \langle S, \mathsf{P}(S), G \setminus \mathsf{P}[S] \rangle$.

Theorem 6.5. *For countable* G :
$Mod(\Gamma^{ok}) = \{K \in sol(\mathsf{G}^{ok}) \mid brd(\mathsf{G}^{ok}) \subseteq \mathsf{P}(K)\} = mPS(\mathsf{G})$.

Proof. ($\supseteq$) follows using Claim 7.3 but we have to show that any $S \in mPS(\mathsf{G})$ actually satisfies Γ. This holds for any G and any $S \in PS(\mathsf{G})$:
(1.a) $S \in PS(\mathsf{G}) \Rightarrow \alpha_S \models \Gamma$. For each $\overline{xy}$ from $x \to y$, we have one of four cases:
(i) $y \in S \Rightarrow x \in \mathsf{P}(S)$
(ii) $x \in S \Rightarrow y \notin S$ and, since $S \in SK(\mathsf{G})$, $y \in \mathsf{P}(S)$
(iii) $x \in \mathsf{P}(S)$ or $y \in \mathsf{P}(S)$
(iv) $\{x, y\} \subseteq \mathsf{G} \setminus \mathsf{P}[S]$
For each $\mathsf{N}[x] = \{x\} \cup Y$, we have one of five cases:
(i) $\exists y \in Y : y \in S$
(ii) $x \in S$
(iii) $x \in \mathsf{P}(S) \to \exists y \in Y : y \in S$
(iv) $\exists y \in Y : y \in \mathsf{P}(S)$, since $S \in PS(\mathsf{G})$, so $x \in \mathsf{P}[S]$, i.e., $\alpha_S \models \mathsf{N}[x]$ by (ii) or (iii)
(v) $\mathsf{N}[x] \subseteq \mathsf{G} \setminus \mathsf{P}[S]$
(1.b) $\forall S \in PS(\mathsf{G}) : \Gamma \vdash \bot(x) \Rightarrow x \notin \mathsf{P}[S]$.

By (1.a) $S \in PS(\mathsf{G}) \Rightarrow \alpha_S \models \Gamma$ so, by soundness Claim 7.3 (with respect to (7.2), for any clause $C : \Gamma \vdash C \Rightarrow \alpha_S \models C$. Hence, $\Gamma \vdash \bot(x) \Rightarrow \alpha_S \models \bot(x)$, i.e., $x \in \mathsf{G} \setminus \mathsf{P}[S]$.
(1.c) Since $\Gamma^{ok} \nvdash \{\}$ by Fact 5.3.2, there is some $K \in sol(\mathsf{G}^{ok})$, i.e., one with $\mathsf{P}[K] = \mathsf{G}^{ok}$. By (2.a), (2.b) below, $K \in PS(\mathsf{G})$, so since $S \in mPS(\mathsf{G})$, we must have $\mathsf{P}[S] \supseteq \mathsf{P}[K] = \mathsf{G}^{ok}$. By (1.b), $x \in \mathsf{P}[S] \Rightarrow x \notin \mathsf{G}^{\bot} \Rightarrow x \in \mathsf{G}^{ok}$, i.e., $\mathsf{P}[S] \subseteq \mathsf{G}^{ok}$, so that $\mathsf{P}[S] = \mathsf{G}^{ok}$. By Eq. 6.2, $\mathsf{P}(\mathsf{G}^{ok}) \subseteq \mathsf{G}^{ok}$, so since G^{ok} is induced subgraph of G, we have the first of the following equalities $\mathsf{P}_{\mathsf{G}^{ok}}(S) = \mathsf{P}(S) = \mathsf{G}^{ok} \setminus S$ – showing that $S \in sol(\mathsf{G}^{ok})$.

If $x \in brd(\mathsf{G}^{ok})$ then $\Gamma \vdash \overline{x}$, so $\alpha_S \models \overline{x}$ and $x \notin S$, i.e., $x \in \mathsf{G}^{ok} \setminus S = \mathsf{P}[S] \setminus S = \mathsf{P}(S)$. Thus $brd(\mathsf{G}^{ok}) \subseteq \mathsf{P}(S)$.

($\subseteq$). (2.a) $Mod(\Gamma^{ok}) \subseteq SK(\mathsf{G})$ is Fact 5.6 from [18].

(2.b) By Eq. 6.2, $\mathsf{N}(\mathsf{G}^{\bot}) \subseteq \mathsf{G}^{\bot}$ so $\mathsf{P}(\mathsf{G} \setminus \mathsf{G}^{\bot}) \subseteq (\mathsf{G} \setminus \mathsf{G}^{\bot})$, i.e., $\mathsf{P}(\mathsf{G}^{ok}) \subseteq \mathsf{G}^{ok}$ and so $\mathsf{P}(\mathsf{P}[\mathsf{G}^{ok}]) = \mathsf{P}(\mathsf{P}(\mathsf{G}^{ok}) \cup \mathsf{G}^{ok}) = \mathsf{P}(\mathsf{P}(\mathsf{G}^{ok})) \cup \mathsf{P}(\mathsf{G}^{ok}) \subseteq \mathsf{P}(\mathsf{G}^{ok}) \cup \mathsf{G}^{ok} = \mathsf{P}[\mathsf{G}^{ok}]$.

(2.c) When $S \in sol(\mathsf{G}^{ok})$ then $\mathsf{P}[S] = \mathsf{G}^{ok}$ and $S \in PS(\mathsf{G})$, by (a) and (b). If $S \notin mPS(\mathsf{G})$, it could be extended to an $S \subset M \in mPS(\mathsf{G})$ (for the infinite case, we apply Zorn's lemma here). This could happen only by adding some elements $E \subseteq \mathsf{G}^{\bot}$ but by (1.b) no such $e \in E$ can belong to any $M \in PS(\mathsf{G})$, since $\Gamma \vdash \bot(e)$. $\square$