Propositional Discourse Logic

Sjur Dyrkolbotn and Michał Walicki
Institute of Informatics
University of Bergen, Norway

Abstract

A novel normal form for propositional theories underlies the logic PDL, which captures some essential features of natural discourse, independent from any particular subject matter and related only to its referential structure. In particular, PDL allows to distinguish vicious circularity from the innocent one, and to reason in the presence of inconsistency using a minimal number of extraneous assumptions, beyond the classical ones. Several, formally equivalent decision problems are identified as potential applications: non-paradoxical character of discourses, admissibility of arguments in argumentation networks, propositional satisfiability, and the existence of kernels of directed graphs. Directed graphs provide the basis for the semantics of PDL and the paper concludes by an overview of relevant graph-theoretical results and their applications in diagnosing paradoxical character of natural discourses.

1   Introduction

The natural discourse, in idealized form, can be seen as a network of cross-references, statements that assert or deny each other. Some statements assert external facts. What should count as fact, however, might be a very contentious issue, in which case it seems safest to regard a fact as nothing more than the statement expressing it. The idealization amounts to abstracting from the specificity of facts, which depend on the actual subject matter, and concentrating on the referential structure, the mutual dependencies between the involved statements. Facts can be then taken as statements which are considered true, independently from any other statements.[1]

A statement is considered true if what it claims is accepted to hold and, typically, all statements of a discourse obtain (at least possible) truth-values corresponding to the status of their claims. Occasionally, however, a discourse malfunctions, resulting in the impossibility of assigning any truth-value to some of its statements. The liar and other standard paradoxes provide obvious examples, but typical situations tend to be more complex. If, for instance, Frank

---

[1] Of course, we are *not* saying that this is what facts *are*, only that they can be treated in this way, without impairing correctness of the formal model. Our model works unchanged also when no such facts are available.

asserts the opposite of John, John asserts the same as Paul and Paul asserts the same as Frank then, at first, the situation may be unclear. But a moment of reflection shows that *regardless* of the actual subject matter, no one can be right and no one can be wrong. If Frank is right, then John is wrong, but then Paul is wrong so Frank must be wrong too. In a similar way all possibilities end up undermining themselves and from this we are forced to conclude, by logic alone, that the discourse has malfunctioned.

We are concerned here with diagnosing the disease, the main symptom of which is that the discourse contains statements that we cannot consistently evaluate as either true or false. The *mere diagnosis* of such cases is of independent interest and importance, because it is only from a proper diagnosis that one can begin to analyze *why* things went awry in a particular case. While the answer to the later question may be specific to the domain and circumstances of the discourse, the question of coherence does not hinge on any such extraneous elements. Certainly, agreement with facts (undisputed statements) is mandatory but, in general, does not suffice for ensuring coherence of the discourse. This problem deserves a separate treatment.

The referential structure of discourses will be represented as directed graphs, which capture many essential properties, circularity in particular, in a simple and intuitively appealing way. The associated logic PDL allows to localize malfunctioning (sub)discourses and gives precise insight into the structural causes of the anomalies, in particular, the vicious circularity. The reported results can be summarized as follows:

Section 2: Deciding if a propositional discourse hides any anomalies or else can be consistently evaluated, has several, formally equivalent decision problems: stable extensions in argumentation networks, the existence of kernels in digraphs and satisfiability of propositional theories. Collecting these equivalences (some known earlier only separately) unifies various fields, simplifying also many proofs.

Section 3: Local kernels (generalizing kernels) of digraphs provide semantics for arbitrary, also inconsistent discourses, and the logic is not explosive, allowing to establish valid consequences also when the discourse is inconsistent, for instance, identifying consistent subdiscourses. Local kernels of a graph G can be captured logically using a simple axiomatization of G in Łukasiewicz's logic Ł3. In particular, we show that kernels correspond to consistent assignments in classical logic while local kernels correspond to consistent assignments in Ł3. We note, however, some shortcomings of Ł3, the central one relating to the difficulties with treating the third value (paradox) in the same was as the two classical ones. Paradox seems to be admitted into a discourse only when it is not possible to find any classical truth assignment. In this sense it has a necessary character, as opposed to the classical values, which must be only possible, for the discourse to be meaningful. In Ł3, as is typical in non-modal logics, questions of possibility (consistency) can only be addressed by some indirect means. We therefore introduce *propositional discourse logic*, PDL. It allows to decide possible truth-values of complex formulae over arbitrary discourse graphs. By

the equivalences from Section 2, it gives the means for deciding:

- paradoxical character of discourses;
- satisfiability of propositional theories;
- acceptability and admissibility of arguments in argumentation networks;
- the existence of (and membership in) local kernels in digraphs.

Meaningful information will be deduced also from inconsistent discourses, but we use only two truth-values with connectives evaluated by the standard rules. Avoiding any extraneous assumptions, which can affect what counts as an anomaly, should make the resulting diagnosis no more dubious than the classical intuitions on which it is based. In light of this, we dare call our logic *essentially* classical.

Section 4: PDL is based on the concept of local kernel and kernel-theory provides valuable results for the analysis of discourses. We cite a series of such results and show their applications in diagnosing problematic cases. At the same time, these results make precise many intuitions, in particular, concerning (vicious) circular reference.

More involved proofs, not included in the text, can be found in the appendix.

Before we embark on the technical parts we present some of our intuitions about natural discourse. Arguing for them might take another paper, so we present them only as a possible motivation, enhancing the understanding of the technical parts. The reader interested only in the latter, can go directly to Section 2.

Elements of natural discourse

Natural discourse is open-ended, it has no marked beginning nor end, there is no period, only "..." preceding and following every statement. Yet, every now and then we have to stop and consider some part of it, some *relative* totality.

Consider the series of consecutive statements, with some longer suspensions, marked by the horizontal lines, at which the possible truth-values of the statements made so far are evaluated, giving the results in the respective column:

a	...The next statement is false...	?	⊥	?	**0**	**0**	⊥
b	The next statement is false...	?	⊥	?	**1**	**1**	⊥
c	The first statement (a) is false ...		⊥				
c'	... and, by the way, so is the next one...			?	**0**	**0**	⊥
d	The next statement is false...			?	**1**	**1**	**0**
e	The previous statement is false...				**0**	**0**	**1**
f	The previous statement is false...				**1**		
f'	... and so is this one...					⊥	**0**

(1.1)

The primed versions (c', f') refer to completion of statements which might have seem completed at the earlier, unprimed, stage. At point b, the truth-values of the two statements are unclear. If the person suspends the voice after c, we may think that he has said the last word, making no sense. But if he continues as

indicated, the discourse becomes again potentially meaningful at c', while at e and f all statements can even be assigned classical truth-values. At f', however, it dissolves again. At this point it exemplifies all phenomena we will address, so there is no need to extend it.

Open-endedness means, in particular, that many statements are *undetermined* at the moment they are made; their truth-value is not exactly known (like, e.g., future contingents). "Snow is white" may be unproblematic, but is a rather special case, representative at most of a special class. The pair $d - e$ illustrates well this modal element. In the absence of any additional information, there seems to be no reason to choose between d and e, and keeping both possibilities is the most natural, not to say ethical, way. Under special circumstances, such sets of possibilities can be narrowed to unique truth-values, resolving indeterminacy into certainty. Yet even with the simplest, empirical claims, one does not have the capacity to verify them all against their eventual justification basis. Most statements are therefore accepted on the basis of other statements, in many situations on the basis of faith, in others on the basis of some defaults or coherence.

Importantly, even when empirical evidence is insufficient, truth-values of many statements can be intuitively ascertained. This does not imply any idealism nor any reduction of truth to coherence, only that it might be difficult to argue with one who does not agree that if e is wrong then d is right. It seems hard to deny that internal coherence is an indispensable feature of meaningful discourse and equally hard to deny that our *approximations* of truth often falter on exactly this point. Most significantly, accepting some statements on the basis of (empirical) facts does not in any way exclude accepting others on the basis of internal coherence. We will show how the two kinds of justification may coexist and how to decide which kind is applicable to a given statement.

The inherent indeterminacy and possible reliance on "mere" coherence reflect the holistic character of such cross-referential networks. Due to mutual dependencies, the discourse can not always be evaluated assigning step by step correct values to single statements. If c' is right, then d must be wrong, but this may, in turn, depend on e, c, etc.[2] Consequently, an anomaly is an accident of *the whole* discourse, not of any particular among its statements. Just like no particular statement among $a - b - c$ is wrong, there is no single culprit among all statements $a - f'$ – they malfunction only *together*. Certainly, easiest to identify are single paradoxical statements, like the liar, but they represent only special cases of discourses, limited to a single statement. There is no need to distinguish such special cases from more complex ones, like $a - b - c$ or $a - f'$, once we accept the anomaly as a holistic phenomenon of the totality of a discourse. The meaning of this, possibly controversial claim, should be transparent in view of the just mentioned examples. It is also in line with more recent developments. In the infinitary Yablo's paradox [36], for instance, no single statement is para-

---

[2]When unfolded in time as a sequence of consecutive statements, such a holistic network of mutual dependencies gives rise to anaphoric and cataphoric references, yielding the non-monotonic character of the discourse. But since non-monotonicity appears thus only as a special, temporal view of mutual dependencies, we will not devote it separate treatment.

doxical, taken on its own. Saying, on the other hand, that every one of them is, requires to consider them in conjunction with all others.

"Holism" does not refer here to any universal totality of everything. We do not deny its possibility but neither know where to find it nor attempt to do it. Locally meaningful, relative totalities, on the other hand, appear every time we conduct a conversation and our holism relies only on such relative totalities of actual interest. Its essential aspect is possible lack of compositionality: a series of meaningful and consistent statements may yield a paradoxical totality. But even from such a totality one can often extract meaningful information, for instance, by focusing on a *subdiscourse* which remains consistent. In fact, what is taken as the *actual* totality bears a crucial influence on the truth-values of the involved statements. If, at point f', we view only the last three statements $d - e - f'$ as the relevant totality, it is consistent. But if we extend it all the way back to a, then there is no way of assigning, in a consistent way, truth-values to all statements – the discourse is paradoxical. Likewise, the discourse which is inconsistent after $a - b - c$, can acquire a promise of potential meaning at c', and even become fully meaningful, allowing classical distribution of truth-values among all its statements, if it ends at e or f.[3]

The inconsistency of a discourse D does not prevent us from deducing useful information about its particular statement x. For instance, if x is true or false in some consistent subdiscourse, then we might want to know – after all, D itself is just a snapshot of some larger totality. Its choice seems, at least in part, guided by the desire to avoid inconsistency. So if we can do better by looking at smaller or larger discourses, why not? A counter-argument might be that it is unclear where to draw the line. Admitting *any* subdiscourse containing x might be too permissive. In discourse (1.1), for instance, stopping after d is hazardous. It refers to the later statement e, so a judgment about d commits us also to a specific judgment about e and such necessary dependencies should be taken into account.

The logic PDL allows us to do this, capturing a natural condition that separates the coherent, acceptable subdiscourses from the others. Loosely expressed, the condition says that *a subdiscourse is an acceptable totality only if one can consistently assign truth-values to its statements so that no extension of this assignment to other statements of the discourse can invalidate it.* Such a complete evaluation of subdiscourse's statements is, if you like, a proof of its admissibility. The condition expresses its robustness – whatever happens to the rest of the discourse does not affect the truth-values within the subdiscourse. "Hard facts", statements accepted as true independently from the rest of the discourse, provide a basic example. Much more involved examples, involving circular and even ungrounded subdiscourses, will be given once we have defined precisely the necessary notions.

---

[3]We do not claim that decision determining the borders of the actually relevant totality is as arbitrary as suggested by the examples. But we do mean that in many cases, the actual totalities have unsharp borders, which may be adjusted in different ways.

2 Formalization

A discourse, over a set of propositional variables Σ, is a finite propositional theory consisting of a series of equivalences

$$x \leftrightarrow \bigwedge_{y \in I_x} \neg y \qquad (2.1)$$

where each $I_x \subseteq \Sigma$ is finite and each $x \in \Sigma$ occurs exactly once on the left of such an equivalence.[4] We use the convention that the right-hand side is **1** when $I_x = \varnothing$. Rendered in this pattern, discourse (1.1) becomes:

$$\begin{array}{llll}
a & \leftrightarrow & \neg b & \qquad d \;\; \leftrightarrow \;\; \neg e \\
b & \leftrightarrow & \neg c' & \qquad e \;\; \leftrightarrow \;\; \neg d \\
c' & \leftrightarrow & \neg a \wedge \neg d & \qquad f' \;\; \leftrightarrow \;\; \neg e \wedge \neg f'
\end{array} \qquad (2.2)$$

The variable on the left of each equivalence acts as the unique identifier of the actually pronounced statement, occurring on its right. The intuitive incoherence of a discourse, the impossibility of assigning truth-values to all its statements (variables on the left), corresponds exactly to the inconsistency of such a theory. Variants of this format were implicit in [9, 22], and elaborated in [34]. It does not limit the expressive power, and provides a normal form for propositional theories, as shown in [5].[5]

The consistency of discourses turns out to be equivalent to two other problems: the existence of stable extensions in argumentation networks and the existence of kernels in digraphs.

2.1 Argumentation networks

Consider the discourse (1.1) with all statements claiming falsity or truth of others replaced by arguments contesting validity of other arguments.

a	...The next argument is wrong...	?	⊥	?	**0**	**0**	⊥
b	The next argument is wrong...	?	⊥	?	**1**	**1**	⊥
c	The first argument (a) is wrong ...	⊥					
c'	... and, by the way, so is the next one...			?	**0**	**0**	⊥
d	The next argument is wrong...			?	**1**	**1**	**0**
e	The previous argument is wrong...				**0**	**0**	**1**
f	The previous argument is wrong...				**1**		
f'	... and so is this one...					⊥	**0**

(2.3)

One seldom encounters arguments like f' in practice, but $a - b - c$ is quite possible.[6] Argumentation theory, at least in its AI version arising from [18],

---

[4]Many results that will be presented hold also for infinite discourses (theories) and infinitary logic (allowing infinite I_x's), but we are addressing primarily the finite and finitary case.

[5]One can think of a propositional letter appearing on the left of an equivalence as naming the complex formula that appears on the right. The equivalences become then instances of Tarski's T-schema, formulated in propositional logic

[6]In [29], p.238, the authors note disappointingly little attention paid to the self-defeating arguments in the argumentation literature. Although psychologically very different from incoherent totalities of arguments, like $a-b-c$, their formal role and effects are entirely analogous.

addresses coherence of such argumentation networks. The analogy between the two examples is obvious, as claims of falsity of other statements act exactly as the arguments attacking other arguments. If a claims falsity of b, and b turns out to be true, then a is false; while if an argument a attacks b and b turns out to be valid/accepted then a becomes defeated/invalidated. We can therefore conflate these two scenarios and represent the fact that a claims falsity of (respectively, attacks) b by an edge in a directed graph $a \rightarrow b$. Discourses (1.1) and (2.3) become thus the graph D': [7]

$$\mathsf{D}' : \qquad \begin{array}{ccc} a & d \rightleftharpoons e \\ \downarrow \nwarrow & \uparrow \quad \uparrow \\ b \longrightarrow c' & f' \, \circlearrowright \end{array} \qquad\qquad (2.4)$$
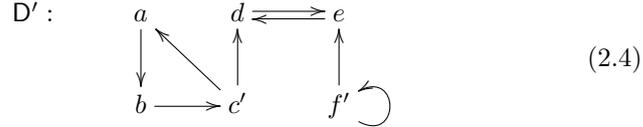
The discourse immediately after c, and before c', is the triangle $a-b-c$ without other nodes nor the edge $\langle c', d \rangle$, while after f, but just before f', is D' without the loop at f', which we will denote D.

Formally, an argumentation network is a directed graph, $\mathsf{G} = \langle G, E \rangle$, with $E : G \rightarrow \mathcal{P}(G)$ determining the out-neighbours of each node. (All functional notation is extended pointwise to sets, e.g., for $X \subseteq G : E(X) = \bigcup_{x \in X} E(x)$. We consider only directed graphs, so "graph" means digraph unless explicitly stated otherwise.) A *solution* is an assignment $\alpha \in \{\mathbf{0}, \mathbf{1}\}^G$ of boolean values $\mathbf{1}$ (true, accepted) or $\mathbf{0}$ (false, defeated), respecting the following rules:

$$\begin{array}{llll} (1) & \forall x \in dom(\alpha) : \alpha(x) = \mathbf{1} & \Leftrightarrow & \forall y \in E(x) : \alpha(y) = \mathbf{0} \\ (2) & \forall x \in dom(\alpha) : \alpha(x) = \mathbf{0} & \Leftrightarrow & \exists y \in E(x) : \alpha(x) = \mathbf{1} \end{array} \qquad (2.5)$$

The rules are equivalent in the context of consistent, classical theories, but we record them both for further use. The set of solutions for a graph is denoted $sol(\mathsf{G})$. In argumentation theory, a solution α corresponds to a *stable extension*, given by the set $\alpha^{\mathbf{1}} = \{x \in G \mid \alpha(x) = \mathbf{1}\}$ ([18], Definition 13, Lemma 14). In terms of discourses, it gives a consistent assignment of truth-values to all statements, i.e., a model of the respective theory (2.1). Non-existence of a solution indicates an anomaly, an incoherent set of arguments or a paradoxical element in the discourse. A trivial example is the liar, an argument defeating itself – the graph $x \, \circlearrowright$ has no solution. A more elaborate example is D' in (2.4). Its lack of coherence can be seen trying, for instance, first to make $d = \mathbf{1}$. This forces $e = \mathbf{0}$ and leaves the liar node f' with its loop without any possible assignment. Trying instead $d = \mathbf{0}$, leaves the triangle $a - b - c$, which can not be assigned any value. D, on the other hand, is not problematic, since assigning

---

[7]Direction of the edges may be reversed, provided that it is done consistently throughout the whole development. Argumentation networks, or various derivative concepts, are typically formulated in the literature with edges going in the opposite direction.

A particular consequence of the representation (2.1) and this graphical counterpart is that statements, the actual carriers of truth-values, correspond to the sentence tokens and not types. Saying the same sentence (type) at two different points may turn it into different statements. A token, or a statement, is in this context just a point in a network of cross-references, a node of the discourse graph.

$d = b = f = \mathbf{1}$ and $a = c' = e = \mathbf{0}$, respects (2.5). Informal verification of the dialogues confirms the intuitive correctness of the conclusion that $sol(\mathsf{D}') = \varnothing$.

Any discourse T, in the form (2.1), gives a graph $\mathcal{G}(\mathsf{T})$, by taking all propositional variables of T as the nodes, and defining the out-neighbours $E(x) = I_x$ for every variable x. In particular, variables which occur on the left-hand side of the equivalences $x \leftrightarrow \mathbf{1}$ become the sinks of $\mathcal{G}(\mathsf{T})$. Conversely, a digraph G gives a discourse $\mathcal{D}(\mathsf{G})$ by taking its nodes G as variables and forming the equivalence $x \leftrightarrow \bigwedge_{y \in E(x)} \neg y$ for each $x \in G$. For the graph D' in (2.4), $\mathcal{D}(\mathsf{D}')$ gives the discourse (2.2), while $\mathcal{G}(\mathcal{D}(\mathsf{D}')) = \mathsf{D}'$. These transformations yield easily the following fact. It only specializes a more general fact from [5], but gives here a sufficient formulation of the equivalence of the logical and graphical versions of the problem. ($mod(X)$ denotes all models of a propositional theory X):

Fact 2.6 *For every graph G and discourse T*
$$sol(\mathsf{G}) = mod(\mathcal{D}(\mathsf{G})), \text{ and}$$
$$mod(\mathsf{T}) = sol(\mathcal{G}(\mathsf{T})).$$

Argumentation theory was worth mentioning both because it is a field of wide interest ([29] gives a good overview, [20] shows newer developments), and because the plain equivalence to the problem of paradox makes the transfer of our results straightforward. But we neither assume familiarity with its details nor intend to present them. We will only parenthetically mention relations to some concepts from argumentation theory. The connections to graphs, on the other hand, are of central importance, as suggested by the above fact and explained further below.

2.2 Kernels of digraphs

A *kernel* of a digraph $\mathsf{G} = \langle G, E \rangle$ is a subset $K \subseteq G$ which is independent (no edges between nodes in K) and absorbing (every node outside K has an edge to some node in K):

$$
\begin{array}{rll}
 & G \setminus K \supseteq E^{\smile}(K) & \text{(independent)} \\
\text{and} & G \setminus K \subseteq E^{\smile}(K) & \text{(absorbing)} \\
\hline
\text{i.e.,} & G \setminus K = E^{\smile}(K), &
\end{array}
\qquad (2.7)
$$

where $E^{\smile}$ denotes the converse of E, i.e., $E^{\smile}(y) = \{x \in G \mid y \in E(x)\}$. One checks easily that K is a kernel iff the assignment $\alpha_K = (K \times \mathbf{1}) \cup ((G \setminus K) \times \mathbf{0})$ is a solution, i.e., satisfies conditions (2.5). For instance, D from (2.4) has a unique kernel, containing nodes assigned $\mathbf{1}$ at point f in (1.1)-(2.3); while D', i.e., D with the additional loop at f', has no kernel, representing paradoxical discourse, the whole $a - f'$.

The main semantic notion associated with our graphical representation, generalizing the notion of a kernel, is a *local kernel*, [28]. It is an independent subset L which absorbs its out-neighbours, i.e., an $L \subseteq G$ satisfying:

$$E(L) \subseteq E^{\smile}(L) \subseteq G \setminus L. \qquad (2.8)$$

One verifies easily that a kernel is a local kernel, while a local kernel need not be a kernel. $Lk(\mathsf{G})$ denotes the set of local kernels in G. In argumentation theory, a local kernel is called an *admissible* extension, and an argument is *acceptable* if it can be added to it, resulting in a new admissible extension. Iterating such a process leads to a *complete* extension $\overline{L}$ – the unique, maximal extension that extends the admissible set L. In terms of graphs, for any local kernel $L \in Lk(\mathsf{G})$ one obtains inductively its *completion*, $\overline{L}$, defined as follows:

Definition 2.9 *The completion* $\overline{L}$ *of an* $L \in Lk(\mathsf{G})$ *is defined inductively:*

$$\begin{aligned} L_0 &= L \\ L_{i+1} &= sinks(\mathsf{G} \setminus E^{\smile}(L_i)) \end{aligned}$$

Fixed-point, $\overline{L} = L_{i+1} = L_i$, *is reached no later than at* $i = |G|$.

For all $i : L_i \in Lk(\mathsf{G})$ and $\mathsf{G} \setminus (\overline{L} \cup E^{\smile}(\overline{L}))$ has no sinks. Of special interest will be the completion of the empty local kernel, $\overline{\varnothing}$, representing the values necessarily induced from the "facts", sinks of the graph. We then let G° be the subgraph of G induced by $G^{\circ} = G \setminus (\overline{\varnothing} \cup E^{\smile}(\overline{\varnothing}))$. It represents the sinkless residuum of G, remaining after removal of all nodes with values induced from the sinks. Since for any $L \in Lk(\mathsf{G}) : sinks(\mathsf{G}) \subseteq sinks(\mathsf{G} \setminus E^{\smile}(L))$, $\overline{\varnothing}$ is obviously contained in the completion of every local kernel:

$$\text{For every } L \in Lk(\mathsf{G}) : \overline{\varnothing} \subseteq \overline{L}. \tag{2.10}$$

Now, for any local kernel $L \in Lk(\mathsf{G})$, the assignment

$$\alpha_L = (L \times \mathbf{1}) \cup (E^{\smile}(L) \times \mathbf{0}) \tag{2.11}$$

is, so to speak, "justified": each node assigned $\mathbf{0}$ has an out-neighbour assigned $\mathbf{1}$, while all out-neighbours of a node assigned $\mathbf{1}$ are assigned $\mathbf{0}$. Interestingly, this is equivalent to satisfaction of (2.5), as ensured by the following fact (recall that for an $\alpha \in \{\mathbf{0}, \mathbf{1}\}^G$, we denote $\alpha^{\mathbf{1}} = \{x \in dom(\alpha) \mid \alpha(x) = \mathbf{1}\}$).

Fact 2.12 *For any graph* G, *subset* $H \subseteq G$ *and* $\alpha \in \{\mathbf{0}, \mathbf{1}\}^H$:
α satisfies both conditions (2.5) iff $\alpha^{\mathbf{1}}$ *is a local kernel of* G *and* $\alpha = \alpha_{\alpha^{\mathbf{1}}}$.

PROOF. The condition (1) implies that $\alpha^{\mathbf{1}}$ must be independent, so $E^{\smile}(\alpha^{\mathbf{1}}) \subseteq G \setminus \alpha^{\mathbf{1}}$ and, moreover, that $E(\alpha^{\mathbf{1}})$ be assigned $\mathbf{0}$. But then (2) requires for any $x \in E(\alpha^{\mathbf{1}})$ to have an edge back to $\alpha^{\mathbf{1}}$, i.e., $E(\alpha^{\mathbf{1}}) \subseteq E^{\smile}(\alpha^{\mathbf{1}})$. The equality $\alpha = \alpha_{\alpha^{\mathbf{1}}}$ is then obvious.

Conversely, making $\alpha_L(L) = \mathbf{1}$ for a local kernel L ensures (1) when also $\alpha_L(E^{\smile}(L)) = \mathbf{0}$. The latter ensures then trivially (2), since for each $x \in E^{\smile}(L)$: $E(x) \cap L \neq \varnothing$. $\qquad\square$

In particular, for a total $\alpha \in \{\mathbf{0}, \mathbf{1}\}^G$, $\alpha^{\mathbf{1}}$ is a kernel of G iff α is its solution which, by Fact 2.6, is equivalent to α being a model of the discourse $\mathcal{D}(\mathsf{G})$.[8]

---

[8]The equivalence of kernels and non-paradoxical discourses was first noted in [9], while of kernels and stable extensions of argumentation networks in [15].

A subdiscourse of $\mathsf{G} = \langle G, E \rangle$ can be seen generally as an arbitrary subgraph of G. More specifically, we will speak about a subdiscourse *induced by a set of statements* $H \subseteq G$, meaning the induced subgraph $\mathsf{H} = \langle H, E \cap (H \times H) \rangle$, or about a subdiscourse *induced by a local kernel* $L \subseteq G$, namely, the subgraph induced by $L \cup E^{\smile}(L)$. The intended meaning should be clear from the context.

Example 2.13 (1) *Sinks of a graph, $sinks(\mathsf{G}) = \{x \in G \mid E(x) = \varnothing\}$, can be seen as "external facts", accepted as true. A statement directly negating such a fact is a node pointing at it. Every collection $L \subseteq sinks(\mathsf{G})$ is a local kernel (since $E(L) = \varnothing$), inducing the assignment of $\mathbf{0}$ to all nodes in $E^{\smile}(L)$.*

(2) *Consider the subdiscourse F of D' from (2.4) induced by $d - e - f'$:*

d : *The next statement is false.*

e : *The previous statement is false.* F: $d \rightleftarrows e \longleftarrow f' \circlearrowright$

f' : *The previous statement is false, and so is this one.*

This subdiscourse arises from the local kernel $E = \{e\}$, as $dom(\alpha_E)$ according to (2.11). The local kernel $\{d\}$ induces even smaller subdiscourse $d \leftrightarrows e$ of F. In either case, the induced assignment respects (2.5) independently *from the values (or their lack) assigned to the rest of D'.*

In terms of discourses, a local kernel gives a consistent – possibly partial – evaluation, which can be seen as internally justified: all its statements can be made simultaneously true, while all statements they claim to be false, are made false. A local kernel L gives thus a general concept of a "coherent subdiscourse", in the sense of a subset of statements, namely $dom(\alpha_L)$, which can be consistently assigned truth-values, obeying the rules (2.5), irrespectively of the assignment to all other statements.[9] For instance, the graph D' from (2.4) has no kernel, but $\{d, b\}$ is its local kernel, and so is $\{e\}$ (the latter inducing the subdiscourse F from Example 2.13.(2).) The lack of any kernel suggests some anomaly, as we can see considering the triangle $a - b - c$ or the whole graph D'. But an anomaly does not mean meaninglessness – the discourse may still possess a lot of information, which can be recovered from its local kernels. These provide the semantic basis for the logic PDL which is introduced in the following section.

3   The Propositional Discourse Logic

The logic PDL allows to establish facts about possible truth or falsity of statements in any finite propositional discourse. Semantics of a discourse is determined by the assignments induced, according to (2.11), from the local kernels of the network of its cross-references, represented by the digraph, as exemplified by (2.4).

---

[9]In argumentation theory, this is referred to as *credulous acceptance*, inroduced in [18], whereby an argument is designated as acceptable when there is a local admissible set containing it. Our work demonstrates that this notion can be looked at as classical satisfiability of a special type of subdiscourse containing the argument, and that both these viewpoints are captured by the technical notion of a local kernel, which has been studied by graph-theorists since the 70ties [28].

Given a graph $\mathsf{G} = \langle G, E \rangle$, we let wff_G be the set of all propositional formulae over the alphabet G, formed using the set of connectives $\{\neg, \wedge\}$.[10] $\Gamma \subseteq \mathsf{wff}_G$ denotes any finite set of such formulae. The *basic* formulae of the language $\mathcal{L}$ have the format $\langle \Gamma : \mathsf{G} \rangle$ and are understood as saying that all formulae in Γ can be made simultaneously true in the discourse (over the graph) G by some (possibly partial) assignment respecting (2.5), i.e., induced by a local kernel according to (2.11). Just as any actual totality of a natural discourse limits the range of possible distributions of truth-values between its statements, so does the graph act as a restriction on the relevant assignments to its nodes, the tokens of the discourse. We see the discourses as *syntactic* objects, a part of the language and not of the semantics. They form the *syntactic contexts* within which claims are made. When reasoning about the truth of these claims, we need to access not only the structure of the claims themselves, but also of the discourse in which they are uttered. This will be further illustrated by the proof system, where inference rules modify the parameter G, along with Γ. (Also for this reasons, the notation $\models \langle \Gamma : \mathsf{G} \rangle$ will be preferred to $\mathsf{G} \models \Gamma$.)

A basic formula $\langle \Gamma : \mathsf{G} \rangle$ is *atomic* if Γ contains only literals. We use the notation $\Gamma^+ = \{a \mid a \in \Gamma\}$ and $\Gamma^- = \{a \mid \neg a \in \Gamma\}$ for the set of positive and negated variables respectively, i.e., such that a basic formula, $\langle \Gamma : \mathsf{G} \rangle$, is atomic if and only if $\Gamma = \Gamma^+ \cup \{\neg a \mid a \in \Gamma^-\}$. The full language $\mathcal{L}$ is given by *composite formulae*, namely, propositional combinations of the basic formulae, again, using only $\neg$ and $\wedge$. Their finite sets Θ, Φ form sequents, $\Theta \vdash \Phi$, using notational conventions of sequent calculi.

Definition 3.1 *The satisfaction relation $\models \subseteq 2^G \times basic(\mathcal{L})$ is defined inductively as follows:*

- $S \models \langle \Gamma : \mathsf{G} \rangle$ *if $\langle \Gamma : \mathsf{G} \rangle$ is atomic and $\Gamma^+ \subseteq S$ and $\Gamma^- \subseteq E^{\smile}(S)$*

- $S \models \langle \Gamma, A \wedge B : \mathsf{G} \rangle$ *iff $S \models \langle \Gamma, A, B : \mathsf{G} \rangle$*

- $S \models \langle \Gamma, \neg(A \wedge B) : \mathsf{G} \rangle$ *iff $S \models \langle \Gamma, \neg A : \mathsf{G} \rangle$ or $S \models \langle \Gamma, \neg B : \mathsf{G} \rangle$*

- $S \models \langle \Gamma, \neg\neg A : \mathsf{G} \rangle$ *iff $S \models \langle \Gamma, A : \mathsf{G} \rangle$*

The true formulae, $\models \subseteq \mathcal{L}$, are now defined relatively to local kernels as follows:

- $\models \langle \Gamma : \mathsf{G} \rangle$ *if there is $L \in Lk(\mathsf{G})$ s.t. $L \models \langle \Gamma : \mathsf{G} \rangle$*

- $\models \neg\phi$ *iff $\not\models \phi$*

- $\models \phi \wedge \theta$ *iff $\models \phi$ and $\models \theta$*

The logical consequence is defined in the standard way, for $\Theta, \Phi \subset \mathcal{L}$:

- $\Theta \models \Phi$ *iff there is $\theta \in \Theta$ such that $\not\models \theta$ or there is $\phi \in \Phi$ such that $\models \phi$.*

---

[10] All other connectives can be defined from $\{\neg, \wedge\}$ in the classical manner. This choice does not in any way limit the expressivity of the language, and is made only for establishing an easy connection to graphs.

It is easy to see that $\models$ is monotone with respect to the first argument – if $L \subseteq M$ and $L \models \langle F : \mathsf{G} \rangle$ then $M \models \langle F : \mathsf{G} \rangle$ for any $F \in \mathsf{wff}_G$.

Note that truth of a formula is defined relatively to a discourse, the $\langle ... : \mathsf{G} \rangle$. Moreover, the definition requires only the *existence* of local kernels, so that $\models \langle \Gamma : \mathsf{G} \rangle$ can be read as "possibly Γ in G".[11] Truth of a basic negation, $\models \neg \langle \Gamma : \mathsf{G} \rangle$, expresses then non-existence of any local kernel satisfying Γ, which can be read as "impossibly Γ in G".

Example 3.2 *The following lists some examples of (un)true statements in the discourse F from Example 2.13.(2). Its only local kernels are $D = \{d\}$ and $K = \{e\}$, where the latter is also a kernel of F.*

1. *$\models \langle d : \mathsf{F} \rangle$ since $d \in D$* $\qquad\qquad$ $\mathsf{F} : \quad d \rightleftharpoons e \longleftarrow f' \circlearrowright$

2. *$\models \langle \neg d : \mathsf{F} \rangle$ since $d \in E^{\smile}(K)$*

3. *$\models \langle d : \mathsf{F} \rangle \wedge \langle \neg d : \mathsf{F} \rangle$ since $\models \langle d : \mathsf{F} \rangle$ and $\models \langle \neg d : \mathsf{F} \rangle$*

4. *$\not\models \langle d \wedge \neg d : \mathsf{F} \rangle$ since for any $L \in Lk(\mathsf{F})$ if $d \in L$ then $d \notin E^{\smile}(L)$*

5. *$\not\models \langle d \wedge e : \mathsf{F} \rangle$ since there is no $L \in Lk(\mathsf{F})$ such that $\{d, e\} \subseteq L$*

6. *$\not\models \langle f' : \mathsf{F} \rangle$ since there is no $L \in Lk(\mathsf{F})$ such that $f' \in L$*

7. *$\models \langle e, \neg f', \neg d : \mathsf{F} \rangle$ since $e \in K$ and $\{f', d\} \in E^{\smile}(K)$*

8. *$\not\models \langle d \wedge \neg (\neg f' \wedge f') : \mathsf{F} \rangle$ since for each $L \in Lk(\mathsf{F}) : d \notin L$ or $f' \notin L \cup E^{\smile}(L)$*

In 7, Γ contains a literal for each variable from F, so this validity means that $\{e\}$ is actually a kernel of the graph F. Validity of 3 and invalidity of 4 corresponds to the non-distributivity of the existential quantifier (or diamond) over conjunction. The former says that there is a local kernel making $d = \mathbf{1}$ and there is one making $d = \mathbf{0}$. The latter claims the existence of a local kernel making both simultaneously. Its justification shows that a contradiction, like $d \wedge \neg d$, is not satisfied in any discourse. But as suggested by 8, also its negation may fail. Such a failure, amounting to the impossibility of assigning either $\mathbf{0}$ or $\mathbf{1}$ to a node, means that the statement does not appear in any acceptable, coherent (sub)discourse. Such statements deserve special attention.

Definition 3.3 *In a graph G, $x \in G$ is a* paradox *iff $\models \neg \langle \neg (\neg x \wedge x) : \mathsf{G} \rangle$*

The definition provides means to move from the meta-level, where paradox is a property – inconsistency – of discourses, to the object-level, where we would like to identify particular statements as paradoxical. Familiar examples turn out as expected. The liar, for instance, must be a paradox, $\models \neg \langle \neg (x \wedge \neg x) : x \circlearrowleft \rangle$, for the simple reason that the graph has no local kernels at all. In the more

---

[11] This generalizes the notion of admissibility of arguments in argumentation theory, which considers only Γ consisting of a single propositional variable.

complex discourse $\mathsf{H} : x_1 \overset{\frown}{\underset{\longleftarrow}{}} x_2 \rightarrow x_3 \rightarrow . \rightarrow s$, $\{s\}$ is (the only) local kernel, and all x_i are paradoxical: $\models \neg\langle\neg(x_i \wedge \neg x_i) : \mathsf{H}\rangle$.

Note that $\langle x \wedge \neg x : \mathsf{G}\rangle$ does not hold in any graph so, in particular, a paradox does not come out here as any dialetheia. According to the above definition, it is a statement which can not possibly witness to the negation of such a contradiction.

The definition captures only statements which are *necessarily* paradoxical, failing to function in *all* acceptable subdiscourses.[12] *Contingent* paradoxes are statements $x \in G$ which are paradoxical only under specific circumstances, expressed by some formula $F \in \mathsf{wff}_G$, i.e., such that:

$$\models \neg\langle F \wedge \neg(x \wedge \neg x) : \mathsf{G}\rangle. \tag{3.4}$$

This validity means the impossibility of satisfying both conjuncts simultaneously in G: whenever F is satisfied, then x *necessarily* becomes paradoxical. In Example 3.2, for instance, 8 confirms the intuition that whenever $d = \mathbf{1}$ in F, then f' becomes paradoxical. To capture the real possibility of x being a paradox, however, the above does not suffice. (3.4) is satisfied, for instance, for any contradiction F. One should, in addition, verify that F indeed can be true, i.e., extend (3.4) with the conjunct expressing the factual possibility of F:

$$\models \langle F : \mathsf{G}\rangle. \tag{3.5}$$

For 3.2.8, for instance, the additional verification of $\models \langle d : \mathsf{F}\rangle$ in 3.2.1, shows that in fact there is an acceptable subdiscourse making f' paradoxical.

Paradox, being a necessary consequence of a discourse, has thus a different status than merely possible truth or falsehood. Trying to bring all three on equal footing would lead to a three-valued logic and involve replacing our existential truth by the universal one. Replacing thus the existential quantifier in the first point of Definition 3.1 by the universal one would not be satisfactory since the empty set is a local kernel in any graph. However, such a move could be made with respect to the completions of local kernels, and we show in the next subsection that this is also achieved by viewing the graph as a theory in Łukasiwicz three-valued logic.

3.1 Łukasiewicz's logic Ł3

We first show that just like the classical models of $\mathcal{D}(\mathsf{G})$ determine kernels of G, so the models of $\mathcal{D}(\mathsf{G})$, viewed now as a theory in Ł3, determine its local kernels. Recall the Ł3-tables for the relevant connectives:

$$
\begin{array}{c|c}
\neg & \\
\hline
\mathbf{1} & \mathbf{0} \\
\mathbf{0} & \mathbf{1} \\
\bot & \bot
\end{array}
\qquad
\begin{array}{c||c|c|c}
\wedge & \mathbf{1} & \bot & \mathbf{0} \\
\hline
\mathbf{1} & \mathbf{1} & \bot & \mathbf{0} \\
\bot & \bot & \bot & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0}
\end{array}
\qquad
\begin{array}{c||c|c|c}
\leftrightarrow & \mathbf{1} & \bot & \mathbf{0} \\
\hline
\mathbf{1} & \mathbf{1} & \bot & \mathbf{0} \\
\bot & \bot & \mathbf{1} & \bot \\
\mathbf{0} & \mathbf{0} & \bot & \mathbf{1}
\end{array}
\tag{3.6}
$$

---

[12] The corresponding idea in Kripke's theory of truth from [27] would be to take as paradox only those sentences which are neither true nor false in any fixed-point. We do not claim that this is appropriate for a general theory of truth, which is not our object.

Because $\leftrightarrow$ occurs only as the main connective forming the equivalences (2.1), their right hand sides are evaluated as in strong Kleene logic, which shares the tables for $\neg$ and $\wedge$ with Ł3. One can therefore introduce there other connectives, disjunction and implication in particular, using classical definitions as in strong Kleene logic. Our restricted use of Łukasiewicz's biconditional captures the difference between treating paradox as a third logical value, where all three appear with the same "necessary character" – and treating it only as a limiting case, the impossibility of classically meeting the intuitive meta-requirement that statements have the same semantic value as the content of what they say.

The semantics of $\leftrightarrow$ in Ł3 captures this identity, leading to the following characterization of local kernels. $\models_L$ denotes satisfaction defined by tables (3.6), with $\mathbf{1}$ as the only designated value. In this context, an assignment α_L induced from a local kernel according to (2.11), is treated as a total 3-valued assignment with $\alpha_L(x) = \bot$ for all $x \notin L \cup E^{\smile}(L)$.

Proposition 3.7 *For a graph G, we have:*
   a) if $L \in Lk(\mathsf{G})$ then $\alpha_{\overline{L}} \models_L \mathcal{D}(\mathsf{G})$, and
   b) for any $\alpha \in \{\mathbf{1}, \mathbf{0}, \bot\}^G$, if $\alpha \models_L \mathcal{D}(\mathsf{G})$ then $\overline{\varnothing} \subseteq \alpha^{\mathbf{1}} \in Lk(\mathsf{G})$.

This allows us to replace local kernels by $\models_L$ in the formulation of the semantics of PDL.

Theorem 3.8 *For $\Gamma \subset \mathsf{wff}_G$:*
   $\models \langle \Gamma : \mathsf{G} \rangle$ *iff there is some $\alpha \in \{\mathbf{1}, \mathbf{0}, \bot\}^G$ such that $\alpha \models_L \mathcal{D}(\mathsf{G})$ and $\alpha \models_L \Gamma$.*

Consequently, any reasoning system for Ł3 can be used to establish validity or contradiction of a formula F in a discourse given by a graph G. The logical consequence $\mathcal{D}(\mathsf{G}) \models_L F$ means that F is true in the completion of *every* local kernel of G. In particular, $\mathcal{D}(\mathsf{G}) \models_L x \leftrightarrow \neg x$ (with Ł3 biconditional) iff $x = \bot$ in every model of $\mathcal{D}(\mathsf{G})$. This is certainly an elegant, logical characterization of necessarily paradoxical statements. Definition 3.3 may be less appealing but, in this respect, PDL coincides with Ł3. One verifies easily that $\models \neg\langle \neg(x \wedge \neg x) : \mathsf{G} \rangle$ iff $\mathcal{D}(\mathsf{G}) \models_L x \leftrightarrow \neg x$. The former states the non-existence of any local kernel of G assigning a truth-value to x, while the latter that every local kernel induces the assignment $x = \bot$.

Our initial intuitions suggested the importance of the undetermined character of typical statements, whose truth is a mere possibility. Unlike Ł3 and most other non-modal logics, PDL captures the possible truth/falsehood as the natural dual to the impossible truth and falsehood of paradoxes. To achieve this in Ł3 one must take an indirect meta-route, for instance, using Theorem 3.8 or the following corollary. The possibility (satisfiability) of a formula F is equivalent to the non-validity of its "negation": it is satisfiable iff $\neg F \vee (F \leftrightarrow \neg F)$ is not valid (with Ł3 biconditional, and $x \vee y = \neg(\neg x \wedge \neg y)$.) Let $\neg \Gamma \vee (\Gamma \leftrightarrow \neg \Gamma)$ denote the disjunction of all respective formulae $\bigvee \{\neg F \vee (F \leftrightarrow \neg F) \mid F \in \Gamma\}$.

Corollary 3.9 *For $\Gamma \subset \mathsf{wff}_G$: $\models \langle \Gamma : \mathsf{G} \rangle$ iff $\mathcal{D}(\mathsf{G}) \not\models_L \neg \Gamma \vee (\Gamma \leftrightarrow \neg \Gamma)$.*

PROOF. $\Rightarrow$) If $\models \langle \Gamma : \mathsf{G} \rangle$ then Theorem 3.8 gives an $\alpha \in \{\mathbf{1}, \mathbf{0}, \bot\}^G$ such that $\alpha \models_L \mathcal{D}(\mathsf{G})$ and $\alpha \models_L \Gamma$. In particular, for every $F \in \Gamma : \alpha \not\models_L \neg F \vee (F \leftrightarrow \neg F)$.
$\Leftarrow$) Since $\mathcal{D}(\mathsf{G}) \not\models_L \neg\Gamma \vee (\Gamma \leftrightarrow \neg\Gamma)$, so there is some $\alpha \in \{\mathbf{0}, \mathbf{1}, \bot\}^G$ such that $\alpha \models_L \mathcal{D}(\mathsf{G})$ and $\alpha \not\models_L \neg\Gamma \vee (\Gamma \leftrightarrow \neg\Gamma)$. That is, for every $F \in \Gamma : \alpha(\neg F) \in \{\mathbf{0}, \bot\}$, i.e., $\alpha(F) \in \{\mathbf{1}, \bot\}$ and $\alpha(F \leftrightarrow \neg F) \in \{\mathbf{0}, \bot\}$. But if $\alpha(F) = \bot$ then $\alpha(F \leftrightarrow \neg F) = \mathbf{1}$, hence $\alpha(F) = \mathbf{1}$ for all $F \in \Gamma$. So $\alpha \models_L \Gamma$, yielding $\models [\Gamma : \mathsf{G}]$ by Theorem 3.8, as desired. $\qquad\square$

For instance, for any cycle X, L3 does not establish anything about the truth of any node $x_i \in X$: $\mathcal{D}(\mathsf{X}) \not\models_L x_i \vee \neg x_i$, since in the absence of sinks, the empty local kernel provides a model of $\mathcal{D}(\mathsf{X})$ with $\bot$ at all nodes. For an odd cycle, e.g., a 3-cycle $\mathsf{A} = a_1 - a_2 - a_3$, L3 establishes the paradoxicality of all three nodes, $\mathcal{D}(\mathsf{A}) \models_L a_i \leftrightarrow \neg a_i$. An even cycle, e.g., a 2-cycle $\mathsf{B} = b_1 \leftrightarrows b_2$, does not satisfy such a formula, $\mathcal{D}(\mathsf{B}) \not\models_L b_i \leftrightarrow \neg b_i$, but there is no L3-consequence of $\mathcal{D}(\mathsf{B})$ which would witness to the absence of paradox. The possibility of any b_i being true follows only by indirect analysis, e.g., using the above corollary $\mathcal{D}(\mathsf{B}) \not\models_L \neg b_1 \vee (b_1 \leftrightarrow \neg b_1) \vee \neg b_2 \vee (b_2 \leftrightarrow \neg b_2)$.

Relevance of such merely possible truth/falsehood has been noted, for instance, in Example 3.2.8. Here, consider the following discourse G:

$$\circlearrowright x \rightleftarrows y \rightleftarrows z \circlearrowleft$$
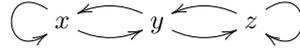
In L3, we have the entailment $\mathcal{D}(\mathsf{G}) \models_L \neg y \to (x \leftrightarrow \neg x)$ (L3 arrows), expressing the intuition that if y is false then x is paradoxical. Although correct, this is not very informative since y cannot possibly be false, which is captured in PDL by $\models \neg\langle\neg y : \mathsf{G}\rangle$. In addition, PDL also gives the possibility of y being true: $\models \langle y : \mathsf{G}\rangle$, i.e., y can be true and can not be false. L3 gives only conditional dependencies, like the one above. In particular, it does not establish the truth of y, since $\bot$ at all nodes is a possible L3-model.

This shortcoming results from the fact that L3 addresses only the semantic information that is already present in the completion $\overline{\varnothing}$ – the truths and falsehoods induced from sinks, the undisputed facts. Although every local kernel provides a model of the discourse, the L3-consequences of a discourse can be determined by looking only at the truths in $\overline{\varnothing}$.

Theorem 3.10 *For $F \in \mathsf{wff}_G$, we have $\mathcal{D}(\mathsf{G}) \models_L F$ iff $\overline{\varnothing} \models \langle F : \mathsf{G}\rangle$.*

PROOF. $\Rightarrow$) For any graph G, $\varnothing$ is a local kernel, so $\alpha_{\overline{\varnothing}} \models_L \mathcal{D}(\mathsf{G})$ by Proposition 3.7.a) and $\alpha_{\overline{\varnothing}} \models_L F$ by assumption. Since F is formed using $\{\neg, \wedge\}$ it is not hard to see, consulting Definition 3.1 and tables (3.6), that $\overline{\varnothing} \models \langle F : \mathsf{G}\rangle$.
$\Leftarrow$) For any α with $\alpha \models_L \mathcal{D}(\mathsf{G})$, $\overline{\varnothing} \subseteq \alpha^{\mathbf{1}}$ by Lemma 3.7.b), so $\alpha^{\mathbf{1}} \models \langle F : \mathsf{G}\rangle$ by the monotonicity of $\models$, and hence $\alpha \models_L F$ (since F is formed using $\{\neg, \wedge\}$). $\square$

As the mere consequences of undisputed facts, these truths from $\overline{\varnothing}$ seem unproblematic. The problematic and more interesting things happen in the sinkless residuum, $\mathsf{G}^\circ = \mathsf{G} \setminus (\overline{\varnothing} \cup E^{\smile}(\overline{\varnothing}))$, as will be further illustrated in Section 4.2. If needed, one can therefore use L3 for analyzing statements which are true in

every acceptable subdiscourse. This novel application of Ł3, although potentially useful, seems however to rest on a concept that is too strong, as witnessed by Theorem 3.10 (so called sceptical semantics in argumentation theory). The inquiry into the alternatives actually present in the acceptable subdiscourses, on the other hand, brings us outside this usual scope of universal truth, away from Ł3 and towards PDL.

3.2 Reasoning in PDL

Consider the 3-cycle, as $a - b - c$ in the introductory example (2.4). Informally, one analyses it by assuming, say, $a = \mathbf{1}$, which requires $b = \mathbf{0}$ and, in turn, $c = \mathbf{1}$. But c can not be true when a is true, so this possibility is excluded. Alternatively, trying $a = \mathbf{0}$ makes $b = \mathbf{1}$ which, in turn, requires $c = \mathbf{0}$. But this, again, gives a conflict since falsity of c means that a is true. In short, and quite generally, we follow the chain of cross-references (to truth of other statements) and assign values, observing the rules (2.5). At the same time, we also decompose the discourse, in the sense that starting with $a = \mathbf{1}$, we never revise this trial but only check "at the end" if the resulting assignment conforms to (2.5). Paradox amounts to the impossibility of assigning either value to some nodes. Informal analysis of the whole discourse G' from (2.4) will be more complex, but along the same lines. Its paradoxical character can be confirmed by observing that the only two possible assignments to $d - e$, make it impossible to assign any values to either the cycle $a - b - c$ or the loop at f'.

The reasoning system PDL, given below in Figure 1, reflects this informal procedure. The composite and basic formula are handled by the standard sequent rules. The non-standard elements are axioms and the first four rules, which address only literals in Γ's. Unlike the standard rules, these decompose both the considered formulae *and* the discourse in which it is evaluated. In this way, and just as the informal analysis sketched above, it is trivially finite and decidable, avoiding any non-terminating revisions of attempted assignments.

A closer look at the rule $(\vdash a)$ should explain the connections to the intuitive procedure above. To establish a possibility of a (being true) in G, $\vdash \langle a : \mathsf{G} \rangle$, the rule's premise requires establishing the possibility of all a's out-neighbours being simultaneously false. This is just (2.5).(1). But the premise is verified in the reduced graph $\mathsf{G} \setminus out(a)$, where $out(a) = \{\langle a, b \rangle \mid b \in E(a)\}$ denotes all edges going out of a. In so reduced graph, a becomes a sink. This is exactly the informal move of assuming a true and checking what happens "at the end", as this value is propagated through the discourse. If such a check "returns to" a without making any of its out-neighbours true, one concludes the possibility of a. This is what happens at the axioms, which require that all things assumed true, end up among the sinks of the resulting, reduced graph.

$\underline{\text{Atomic formulae}}$ (literals $a, \neg a$ in Γ):

Axioms: $\langle \Gamma, \neg a : \mathsf{G} \rangle, \Theta \vdash \Phi$ if $a \in sinks(\mathsf{G})$ $\qquad \Theta \vdash \Phi, \langle \Gamma : \mathsf{G} \rangle$ if $\Gamma \subseteq sinks(\mathsf{G})$

$$(a\vdash) \quad \frac{\langle \Gamma \cup \{\neg a_i \mid a_i \in E(a)\} : \mathsf{G} \setminus out(a) \rangle, \Theta \vdash \Phi}{\langle \Gamma, a : \mathsf{G} \rangle, \Theta \vdash \Phi} \qquad \text{if } E(a) \neq \varnothing$$

$$(\vdash a) \quad \frac{\Theta \vdash \Phi, \langle \Gamma \cup \{\neg a_i \mid a_i \in E(a)\} : \mathsf{G} \setminus out(a) \rangle}{\Theta \vdash \Phi, \langle \Gamma, a : \mathsf{G} \rangle}$$

$$(\neg\vdash) \quad \frac{\langle \Gamma, a_1 : \mathsf{G} \rangle, \Theta \vdash \Phi \ ; \ ... \ ; \ \langle \Gamma, a_n : \mathsf{G} \rangle, \Theta \vdash \Phi}{\langle \Gamma, \neg a : \mathsf{G} \rangle, \Theta \vdash \Phi} \quad \text{if } \{a_1, ..., a_n\} = E(a) \neq \varnothing$$

$$(\vdash\neg) \quad \frac{\Theta \vdash \Phi, \langle \Gamma, a_1 : \mathsf{G} \rangle, ..., \langle \Gamma, a_n : \mathsf{G} \rangle}{\Theta \vdash \Phi, \langle \Gamma, \neg a : \mathsf{G} \rangle}$$

The side conditions on $E(a)$ apply to both rules between which they appear.

$\underline{\text{Basic formulae}}$ (one-sided sequent rules):

$$(\wedge\vdash) \ \frac{...\langle \Gamma, A, B : \mathsf{G} \rangle \vdash ...}{...\langle \Gamma, A \wedge B : \mathsf{G} \rangle \vdash ...} \qquad\qquad (\vdash\wedge) \ \frac{... \vdash \langle \Gamma, A, B : \mathsf{G} \rangle, ...}{... \vdash \langle \Gamma, A \wedge B : \mathsf{G} \rangle, ...}$$

$$(\neg\neg\vdash) \ \frac{...\langle \Gamma, A : \mathsf{G} \rangle \vdash ...}{...\langle \Gamma, \neg\neg A : \mathsf{G} \rangle \vdash ...} \qquad\qquad (\vdash\neg\neg) \ \frac{... \vdash \langle \Gamma, A : \mathsf{G} \rangle, ...}{... \vdash \langle \Gamma, \neg\neg A : \mathsf{G} \rangle, ...}$$

$$(\neg\wedge\vdash) \ \frac{...\langle \Gamma, \neg A : \mathsf{G} \rangle \vdash ... \ ; \ ...\langle \Gamma, \neg B : \mathsf{G} \rangle \vdash ...}{...\langle \Gamma, \neg(A \wedge B) : \mathsf{G} \rangle \vdash ...}$$

$$(\vdash\neg\wedge) \ \frac{... \vdash \langle \Gamma, \neg A : \mathsf{G} \rangle, \langle \Gamma, \neg B : \mathsf{G} \rangle, ...}{... \vdash \langle \Gamma, \neg(A \wedge B) : \mathsf{G} \rangle, ...}$$

$\underline{\text{Composite formulae}}$ (two-sided sequent rules):

$$((\neg\vdash)) \ \frac{\Theta \vdash \Phi, \phi}{\neg\phi, \Theta \vdash \Phi} \qquad\qquad ((\vdash\neg)) \ \frac{\phi, \Theta \vdash \Phi}{\Theta \vdash \Phi, \neg\phi}$$

$$((\wedge\vdash)) \ \frac{\phi, \psi, \Theta \vdash \Phi}{\phi \wedge \psi, \Theta \vdash \Phi} \qquad\qquad ((\vdash\wedge)) \ \frac{\Theta \vdash \Phi, \phi \ ; \ \Theta \vdash \Phi, \psi}{\Theta \vdash \Phi, \phi \wedge \psi}$$

Figure 1: The reasoning system PDL.

Example 3.11 *In a 4-cycle, $a - b - c - d$, a (or any other node) can be true:*

$$\frac{\dfrac{\dfrac{\dfrac{\dfrac{a \in sinks(\ b \twoheadrightarrow c \quad d \twoheadrightarrow a\ )}{\vdash \langle a : \ b \twoheadrightarrow c \quad d \twoheadrightarrow a\ \rangle}}{\vdash \langle \neg d : \ b \twoheadrightarrow c \quad d \twoheadrightarrow a\ \rangle} (\vdash\neg)}{\vdash \langle c : \ b \twoheadrightarrow c \twoheadrightarrow d \twoheadrightarrow a\ \rangle} (\vdash c)}{\vdash \langle \neg b : \ b \twoheadrightarrow c \twoheadrightarrow d \twoheadrightarrow a\ \rangle} (\vdash\neg)}{\vdash \langle a : \ b \rightleftarrows c \twoheadrightarrow d \twoheadrightarrow a\ \rangle} (\vdash a)$$

17

Proofs of paradoxicality use the same graph reductions, but involve two sub-proofs, showing the impossibility of being true and of being false.

Example 3.12 *x with the liar loop is paradoxical – it can be neither true (the left branch) nor false (the right branch):* [13]

$$
(\neg\neg\,\vdash)\,\cfrac{(x\vdash)\,\cfrac{x \in sinks(x)}{\cfrac{\langle \neg x : x\rangle \vdash}{\langle x : \langle x,x\rangle\rangle \vdash}}}{\cfrac{\langle \neg\neg x : \langle x,x\rangle\rangle \vdash \qquad (x\vdash)\,\cfrac{x \in sinks(x)}{\cfrac{\langle \neg x : x\rangle \vdash}{[\langle \neg x : \langle x,x\rangle\rangle \vdash}}\,(\neg\vdash)}{\cfrac{\langle \neg(\neg x \wedge x) : \langle x,x\rangle\rangle \vdash}{\vdash \neg\langle\neg(\neg x \wedge x) : \langle x,x\rangle\rangle}\,((\vdash\neg))}\,(\vdash \neg\wedge)}
$$

The following proof shows that a, in the triangle a-b-c, is impossibly true (left branch) and impossibly false (right branch; the same holds also for b and c):

$$
((\vdash\neg))\,\cfrac{(a\vdash)\,\cfrac{(\neg\vdash)\,\cfrac{(c\vdash)\,\cfrac{a \in sinks(a,c,\langle b,c\rangle)}{\cfrac{\langle \neg a : a, c, \langle b,c\rangle\rangle \vdash}{\langle c : a, \langle b,c\rangle, \langle c,a\rangle\rangle \vdash}}}{\langle \neg b : a, \langle b,c\rangle, \langle c,a\rangle\rangle \vdash}}{\langle a : \langle a,b\rangle, \langle b,c\rangle, \langle c,a\rangle\rangle \vdash}}{\vdash \neg\langle a : \langle a,b\rangle, \langle b,c\rangle, \langle c,a\rangle\rangle} \qquad ((\vdash\neg))\,\cfrac{(\neg\vdash)\,\cfrac{(b\vdash)\,\cfrac{(\neg\vdash)\,\cfrac{(a\vdash)\,\cfrac{b \in sinks(b,a,\langle c,a\rangle)}{\cfrac{\langle \neg b : b, a, \langle c,a\rangle\rangle \vdash}{\langle a : b, \langle a,b\rangle, \langle c,a\rangle\rangle \vdash}}}{\langle \neg c : b, \langle a,b\rangle, \langle c,a\rangle\rangle \vdash}}{\langle b : \langle a,b\rangle, \langle b,c\rangle, \langle c,a\rangle\rangle \vdash}}{\langle \neg a : \langle a,b\rangle, \langle b,c\rangle, \langle c,a\rangle\rangle \vdash}}{\vdash \neg\langle \neg a : \langle a,b\rangle, \langle b,c\rangle, \langle c,a\rangle\rangle}
$$

$$
\cfrac{\vdash \neg\langle a : \langle a,b\rangle, \langle b,c\rangle, \langle c,a\rangle\rangle \quad \wedge \quad \neg\langle \neg a : \langle a,b\rangle, \langle b,c\rangle, \langle c,a\rangle\rangle}{}\,((\vdash\wedge))
$$

The second proof suggests that there may be other characterizations of a paradox, besides Definition 3.3. Indeed, inspecting the rules, we see the equivalence of the provability of the following formulae:

$$\vdash \neg\langle\neg(x \wedge \neg x) : \mathsf{G}\rangle \quad \Leftrightarrow \quad \vdash \neg\langle\neg x : \mathsf{G}\rangle \wedge \neg\langle x : \mathsf{G}\rangle. \tag{3.13}$$

Hence, the second proof in Example 3.12 shows that a is paradoxical according to Definition 3.3 (assuming soundness of PDL, which is proven in the appendix).

This is a special case of the general equivalence, corresponding to the distributivity of universal quantifier (or box) over conjunction:

$$\vdash \neg\langle\neg(A \wedge B) : \mathsf{G}\rangle \quad \Leftrightarrow \quad \vdash \neg\langle\neg A : \mathsf{G}\rangle \wedge \neg\langle\neg B : \mathsf{G}\rangle.$$

As noted before, we have thus a logic of *possible truth and falsehood*, in the context where *tertium datur*, namely, the paradox. However, a paradoxical statement is not a functional consequence of the contingent values assigned to its substatements. It is not a mere possibility but occurs only as the impossibility of truth and of falsehood. It is always necessary and appears only as an unavoidable consequence of the discourse.

---

[13]For displaying proofs, it is convenient to write a graph as a list of sinks and edges, e.g., $\langle x,x\rangle$ is the liar graph, while x the same graph with the loop removed. Some redundancy in notation may ease readability, e.g., $a, \langle b,a\rangle$ and $\langle b,a\rangle$ denote the same graph $b \to a$.

An interesting phenomenon is that when paradox, as a property of statements, is understood in this way, then a discourse can be paradoxical without having any paradoxical statements. It may be namely undetermined which part of the discourse is paradoxical. In the discourse (2.4), depending on the choice of the local kernel for $d-e$, either $a-b-c'$ becomes paradoxical, or else only f'. In this way, a network of contingent paradoxes, captured using pattern (3.4), might be, as a whole, a necessary paradox. Holism strikes back, as it where, and rightly so, since our judgment about the paradoxicality of particular statements is *derived* from the discourse. Traditionally simple examples, like the liar, do not contradict this in any way. They provide only examples of very simple discourses, but not any argument for restricting paradox to single statements.

Provability of the following conditional paradoxes in D' from (2.4) is left as a simple exercise to the interested reader:

$\vdash \neg\langle d \wedge \neg(f' \wedge \neg f') : \mathsf{F}'\rangle$, i.e., when $d = \mathbf{1}$ then f' is paradoxical, and

$\vdash \neg\langle \neg d \wedge \neg(c' \wedge \neg c') : \mathsf{F}'\rangle$, i.e., when $d = \mathbf{0}$ then c' is paradoxical.

PDL is sound and complete with respect to the semantics from Definition 3.1. (The proof is in the appendix. Inspecting the rules, in particular, for literals in Γ, one verifies easily decidability of PDL.)

Theorem 3.14 *For all finite sets* $\Theta, \Phi \subset \mathcal{L} : \Theta \models \Phi \Leftrightarrow \Theta \vdash \Phi$.

3.3   A structure of paradox

PDL provides a tool for detailed investigation of the paradoxical character of particular discourses but, as we saw in Section 3.1, also Ł3 could be used for this purpose. PDL's ability to handle merely possible truth was mentioned as its advantage over Ł3. Another advantage is PDL's sequent calculus. Analysis of the proofs of paradoxicality provides an insight into the general structure of paradox as we will now show.[14] Let's keep in mind here that out-neighbours of a node x represent the statements directly negated by x.

The equivalence (3.13) gives a simple example stating that a node necessarily violates the law of excluded middle if and only if it can not possibly be made false and can not possibly be made true. This is hardly unexpected, but correct and simple expression of basic intuitions is as reassuring as it may be non-trivial.

More interestingly, the fact that x is paradoxical, i.e., impossibly true and impossibly false, can be expressed equivalently as *both x and all its out-neighbours being impossibly true*. Indeed, impossibility of assigning $\mathbf{0}$ to a node amounts to the non-existence of any local kernel containing some of its out-neighbours, and provability in PDL satisfies the equivalence

$$\vdash \neg\langle x : \mathsf{G}\rangle \wedge \neg\langle \neg x : \mathsf{G}\rangle \iff \Big(\big(\vdash \neg\langle x : \mathsf{G}\rangle\big) \text{ and } \big(\vdash \neg\langle y_i : \mathsf{G}\rangle \text{ for all } y_i \in E(x)\big)\Big).$$

---

[14]It is not clear whether the sequent calculus for Ł3 presented in [6] could be used in a similar way. This seems rather unlikely, in particular, as it has multiple rules with the same principal formula.

This follows immediately from the rule $(\neg\vdash)$, which is the only one yielding the impossibility of x being false, $\langle\neg x : \mathsf{G}\rangle \vdash$, and requiring for this the premises stating impossibility of y_i being true, $\langle y_i : \mathsf{G}\rangle \vdash$, for every $y_i \in E(x)$:

$$((\vdash\neg)) \;\; \cfrac{\cfrac{\langle x : \mathsf{G}\rangle \vdash}{\vdash \neg\langle x : \mathsf{G}\rangle}}{\vdash \neg\langle x : \mathsf{G}\rangle \wedge \neg\langle\neg x : \mathsf{G}\rangle} \quad \cfrac{\cfrac{\cfrac{\langle y_1 : \mathsf{G}\rangle \vdash ; \;...\; ; \; \langle y_n : \mathsf{G}\rangle \vdash}{\langle\neg x : \mathsf{G}\rangle \vdash} \;(\neg\vdash)}{\vdash \neg\langle\neg x : \mathsf{G}\rangle} \;((\vdash\neg))}{}$$
$$((\vdash\wedge))$$

Another characterization of paradox can be read from further analysis of the proof of the first conjunct. It is established according to the rule $(x\vdash)$, only when it is not possible for all of x's out-neighbours to be false simultanously. Let $F \subseteq E(x)$ be a maximal subset of $E(x)$ such that $\vdash \langle\bigwedge_{y\in F} \neg y : \mathsf{G}\rangle$. Then the two subproofs together mean that if x is paradoxical then it has an out-neighbour y_i that is paradoxical contingent on $\bigwedge_{y\in F} \neg y$:

$$\vdash\neg\langle\neg(x \wedge \neg x) : \mathsf{G}\rangle \implies \vdash\neg\langle \bigwedge_{y\in F} \neg y \wedge \neg(\neg y_i \wedge y_i) : \mathsf{G}\rangle \text{ for some } y_i \in E(x) \setminus F.$$

(3.15)

The implication can not be reversed, as illustrated by the following example: $\overset{\curvearrowleft}{\bigcirc} y_1 \;\leftarrow\; x \rightarrow y_2$. Although x has a paradoxical out-neighbour y_1, it is not itself paradoxical since the sink y_2 is a local kernel, making $x = \mathbf{0}$. What does hold, however, is that if all out-neighbors of x are impossibly true and at least one of them is also impossibly false, i.e., paradoxical, then x itself is paradoxical. In the following discourse P, x has a contingently paradoxical out-neighbour y_1, but for this to make x paradoxical, also y_2 can not be true.
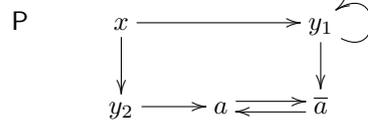
When it comes to contingent paradoxicality, the general schema of the proof of paradoxicality of x, with $E(x) = \{y_1, ..., y_n\}$, shows the sufficiency and necessity of the following two assumptions:

$$(x\vdash)\;\cfrac{(\neg\neg\vdash)\;\cfrac{(\neg\wedge\vdash)\;\cfrac{\cfrac{\langle F, \neg y_1, ..., \neg y_n : \mathsf{G} \setminus out(x)\rangle \vdash}{\langle F, x : \mathsf{G}\rangle \vdash}}{\langle F, \neg\neg x : \mathsf{G}\rangle \vdash}}{((\vdash\neg))\;\cfrac{\langle F, \neg(\neg x \wedge x) : y\mathsf{G}\rangle \vdash}{\vdash \neg\langle F, \neg(\neg x \wedge x) : \mathsf{G}\rangle}} \quad \cfrac{\langle F, y_1 : \mathsf{G}\rangle \vdash ; ...; \langle F, y_n : \mathsf{G}\rangle \vdash}{\langle F, \neg x : \mathsf{G}\rangle \vdash}\;(\neg\vdash)}$$

(3.16)

The premise in the right branch demands the impossibility of any y_i being true along with F, while the left demands the impossibility of all y_i being simultaneously false.

In P, both y_1 can be false and so can y_2. In particular, y_1 is not necessarily paradoxical. But they can not be false simultaneously. When $a = \mathbf{1}$, then x is

paradoxical, but when $a = \mathbf{0}$ then the graph has a kernel, i.e.: $\vdash \neg\langle a, \neg(x \wedge \neg x) : \mathsf{P}\rangle$ and $\vdash \langle \overline{a}, y_2 : \mathsf{P}\rangle$. The latter says that $\{\overline{a}, y_2\}$ is a local kernel of P and it can be extended to the provable claim $\vdash \langle \overline{a}, y_2, \neg a, \neg x, \neg y_1 : \mathsf{P}\rangle$. Since each node of P figures in one literal of the claim, this shows that $\{\overline{a}, y_2\}$ is actually a kernel of P.

Implication (3.15) shows that circularity is indispensable for obtaining a finitary paradox. A paradox must negate a (contingent) paradox.[15] In case of the liar, such a paradoxical out-neighbour is the liar itself while in general, it requires its own paradoxical out-neighbour, etc.. Hence, in a finite graph, a paradox requires a cycle. (The only alternative would be an infinite chain of paradoxical statements, but this requires moving to infinite and, as we will see in the next section, infinitary discourses.) Although this has always been a basic intuition about (finitary) paradoxes, we are not aware of any other, general and strictly formal expression of this idea.[16]

Instead of continuing this logical analysis, we switch now to a graphical one. It captures evil circularity in a more direct and more precise way, providing also a series of general results useful for the diagnosis of discursive anomalies.

4 Some applications of kernel theory

The kernel-theoretic approach provides new means and several results for the analysis of discourses which are often easier and more intuitive than those offered by classical logic. It informs and extends accepted intuitions in a formally precise, yet intuitively appealing way. This section illustrates applicability of kernel theorems for diagnosing paradoxical character of discourses.

Let us only observe that, as seen so far, the truth-operator (or truth-predicate) plays no essential role for the appearance of paradoxes. Having represented them as inconsistent discourses, the object-level negation suffices. This may not reflect all intentions and intuitions about the truth-operator, but does not affect the correctness of the diagnosis of (non)paradoxicality. Questions about the nature of the truth-operator are different than the questions about paradox and the latter can be addressed fruitfully without settling the former. With this reservation, all particular claims about (non)paradoxical character of specific,

---

[15]Curry's paradox may be negation-free only if $x \leftrightarrow (x \rightarrow y)$ does *not* abbreviate $x \leftrightarrow \neg(x \wedge \neg y)$. In our case, it is exactly what it does, as the arrow $\rightarrow$ on the right is defined as in strong Kleene logic.

[16]This may require a qualification. On the one hand, purely logical means are inherently inadequate, since the language of classical logic is designed exactly so as to prevent any direct self-reference. Typically, one is forced to step beyond first-order logic and apply intricate Gödelizations in order to express something as simple as the liar (as a matter of fact, only something which merely reminds of the liar). On the other hand, one may take a more semantic approach. A good example is the use of non-well-founded sets, that is, eventually arbitrary graphs in [3], as the semantic basis for modeling circularity of discourses. Accepting the anti-foundation axiom is, however, a dramatic step, bringing us out of classical set theory. It may happen that a general solution to paradoxes of all kinds might require such a fundamental departure. We prefer to avoid it as long as possible, in particular, when it suffices to represent circularity in classical set theory and analyze it using essentially classical logic.

finitary discourses in the following examples can be proven in PDL. But the presentation should benefit from dispensing with such detailed formalities and keeping it at a more intuitive level.

Occasionally, we address also infinite cases, possibly even in infinitary logic (admitting infinite conjunctions in equivalences (2.1)). Although PDL would require extensions (to infinitary formulae and rules), the basic semantic facts hold unchanged: transformations $\mathcal{D}$, $\mathcal{G}$ and Fact 2.6, definitions in Section 2.2 with Fact 2.12 – all these retain their validity when passing to infinitary logic. Graphically, infinitary logic corresponds to digraphs with infinite branching, and the main difference is that such discourses, unlike the finitary ones, may be inconsistent without involving any circularity.

4.1 Circularity

Circularity seems inherently difficult to capture by logical means alone. Some cases are vicious, others are not and although it has always seemed the key to paradox, not only its nature but even its very occurrence may be disputed. The amount of implicit agreement, underlying most of its discussions, fails in the face of more complex examples or, perhaps, of more involved notions of circularity. For instance, although Yablo's paradox appears at first sight uncontroversially non-circular, this has been challenged and disputed by a series of authors, e.g., [30, 32, 4, 10], some claiming it to possess a sort of circularity. One can construe circularity so that it applies to Yablo's paradox, but this is then a different notion from the simple one, which does not apply to it. The graphical representation offers the standard notion of a cycle which is hardly disputable. A finite path in a graph is a sequence of nodes $x_0 x_1 x_2 ... x_n$, where for all $0 \leq i < n : x_{i+1} \in E(x_i)$. A path is simple when it has no repeating vertices. A cycle is a path $x_0 x_1 x_2 ... x_n$, which is simple except for $x_n = x_0$. The cycle is odd/even when n is. A special case is an odd cycle of length 1, i.e., a loop xx, when $x \in E(x)$.[17]

4.1.1 Only cycles are vicious

According to the theorem from [33], which appears to be the first result in kernel theory, *every finite, directed acyclic graph (dag) has a unique kernel*. By the equivalence with the propositional theories and the compactness theorem, this

---

[17]This excludes any "infinite cycle" and makes Yablo's paradox non-circular. (Infinite cycles can be introduced into infinite graphs, by topological means, using completions of infinite rays. They seem to have no relation to infinitary paradoxes, though, and Yablo remains acyclic also when such cycles are allowed.) Yablo's circularity, suggested in [30], concerned its finitary formulation but not its actual referential structure. This is the relevant structure, captured by our graphs. On the other hand, since every person in the Yablo's path says (*) "All my followers are lying", Priests suggests that "one individuates the thought in such a way that all the people are thinking the same thought". This is certainly possible, but asks us to ignore the crucial structure of the reference involved: the thought of the n-th person includes the $(n+1)$-th person, while the thought of the $(n+1)$-th person does not. As observed in footnote 7, one can plausibly ask also here to individuate the thought (*) – if one wants to insist on the singular form – at the level of tokens and not of its type. The isomorphism of every tail of the Yablo graph with the whole graph does not mean that they are identical.

extends to finitely branching, infinite dags, [5]. As long as statements refer only to finitely many other statements, paradoxicality will only arise from circularity, and this holds even when one considers infinitely many statements. The infinite path of statements $x_0 x_1 x_2 x_3...$, each saying: "The next 3 statements are false." is not paradoxical. Its theory, for all $i \in \mathbb{N} : x_i \leftrightarrow \neg x_{i+1} \wedge \neg x_{i+2} \wedge \neg x_{i+3}$, is consistent, corresponding to a finitely branching dag. No matter what the statements say, as long as each claims something only about finitely many of its followers, the discourse is not paradoxical.

Thus, only cycles can become vicious in finitary discourses and examples abound. A chordless odd cycle has no kernel. This obvious fact subsumes the simplest paradoxes. The liar, $x \leftrightarrow \neg x$, is a loop, $x \,\begin{smallmatrix}\curvearrowright\end{smallmatrix}$, and "I am not true" or "I am not non-false", $x \leftrightarrow \neg\neg\neg x$, is a 3-cycle, $x \rightleftarrows y \longrightarrow z$.

A more general statement is that *a non-empty, finite, sinkless graph, which has no even cycle has no kernels*, [35]. For instance, an odd number n of persons standing in a ring, with every x_i claiming that his successor x_{i+1} is lying and so does the predecessor of his predecessor, x_{i-2} (with addition and subtraction modulo $n - 1$), forms a paradox. For $n = 3$ this is just a 3-cycle, but for larger n this involves chords, as shown for such a paradoxical discourse with $n = 7$: [18]
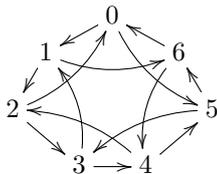
4.1.2 Vicious cycles are odd

Although circularity is necessary for finitary paradoxes, it remains innocent as long as it does not result in any self-negation. The standard example is the truth-teller, which can be formulated in different ways, all giving the same graphical representation:

(1) "This statement is true." or

(2) "This statement is not false." or

(3) "The next statement is false." and "The previous statement is false."

The corresponding theory – $x \leftrightarrow \neg \overline{x}$ and $\overline{x} \leftrightarrow \neg x$ – gives a 2-cycle $x \rightleftarrows \overline{x}$ with two solutions, each assigning complementary values to both statements. [19]

---

[18]Incidentally, this form of discourse (a ring of size $n \geq 3$ where each x_i claims falsity of x_{i+1} and x_{i-2}) is paradoxical even when the ring is even, but this follows from a particular argument concerning the impossibility of breaking the involved odd 3-cycles. (To see this, assume a solution, pick a node x_i that is **1** and look for any **1**-successor of its **0**-successor x_{i+1} on the ring. No such can exist, since all successors of x_{i+1} are in- or out-neighbours of x_i.)

[19]One can propose finer criteria for distinguishing statements, so that (1)-(3) come out as different, even to the point where (3) becomes a No-No paradox. But as far as their truth-conditions under the classical semantics are concerned, there is as little problem with their equivalence – and the absence of paradox – as with the fact that among two persons accusing each other of lying, only one is telling the truth, the symmetry of appearances notwithstanding. Which one it is, may vary between various tokens of truth-teller.

The informal intuition that circularity may be vicious only when it involves some sort of self-negation, is captured precisely by the central result of kernel theory, Richardson's theorem from the early 50-ties, [31], stating that *every finitely branching graph with no odd cycles has a kernel.* Solvability of finitely branching dags is its special case. As a more complex example, consider the infinite T with the statements, for all integers $i \in \mathbb{Z}$, of the form:

$$
\begin{aligned}
x_{2i} &\leftrightarrow \neg x_{2i-1} \wedge \neg x_{2i+1} \\
x_{2i+1} &\leftrightarrow \neg y_{2i+1} \\
y_{2i} &\leftrightarrow \neg x_{2i} \\
y_{2i+1} &\leftrightarrow \neg y_{2i} \wedge \neg y_{2i+2}
\end{aligned}
$$

Its finitely branching graph $\mathcal{G}(\mathsf{T})$ has the form

$$
\begin{array}{ccccccccc}
\ldots & \leftarrow & y_1 & \rightarrow & y_2 & \leftarrow & y_3 & \rightarrow & y_4 & \leftarrow & \ldots \\
 & & \uparrow & & \downarrow & & \uparrow & & \downarrow & & \\
\ldots & \rightarrow & x_1 & \leftarrow & x_2 & \rightarrow & x_3 & \leftarrow & x_4 & \rightarrow & \ldots
\end{array}
$$

Since $\mathcal{G}(\mathsf{T})$ has no odd cycles, T is not paradoxical.

4.1.3 Not all odd cycles are vicious

Richardson's theorem has been generalized in various ways by giving conditions on the odd cycles ensuring the existence of a kernel. For instance, *a finite G has a kernel if each of its odd cycles $C = x_0 x_1 \ldots x_{2k+1}$ has at least:*

  a) *two reversed edges ($x_{i+1} \in E(x_i)$ is reversed if also $x_i \in E(x_{i+1})$), [16],*

  b) *two crossing consecutive chords, [17], or*

  c) *two chords whose targets are two consecutive nodes of the cycle, [24].*

Five persons in a ring, each accusing his right neighbour of lying, form an odd cycle and a paradoxical discourse. By a), if two persons accuse, in addition, also the person to their left, the paradox is resolved. Incidentally, for an isolated odd cycle it is sufficient for only one person to make such an additional claim, but the general result, for arbitrary finite graphs, requires two.

The conditions become more complex as one tries to cover more cases left open by the elegant theorem of Richardson ([7] lists some more results.) As in the case of finite satisfiability, the intractability of the problem of kernel existence, [8, 11], leaves little hope for any compositional criteria for deciding if odd cycles in a given discourse are vicious or not.

4.2 Ungroundedness

Ungroundedness, as introduced by Kripke in [27], subsumes circularity and relates to the issue of contingency. According to Kripke's terminology a statement x is grounded, modulo some monotone operator on partial semantic assignments, if starting from some collection of true atomic statements, the truth/falsity of x is determined by the semantic assignment that is the least fixed-point of the operator in question. (The typical operator used is obtained from the inductive

step in the definition of satisfaction in strong Kleene logic.) The construction has a natural expression in terms of graphs. According to definition (2.7), sinks belong to every kernel, encoding true atomic statements, $x \leftrightarrow \mathbf{1}$. Now, as sinks belong to every kernel, their predecessors do not belong to any, so we can designate all sinks true and all predecessors of sinks false. Iterating this process leads to the assignment $\alpha_{\overline{\varnothing}}$, obtained as in (2.11), where $\overline{\varnothing}$ is the completion of the trivial local kernel $\varnothing$, cf. Definition 2.9. In argumentation theory, the induced assignment $\alpha_{\overline{\varnothing}}$ gives the so called sceptical (or grounded) semantics. The following fact, originating from [31] and stated generally in [5], gives substance to our claim that, for the general purposes, it is of very limited value. Paradoxical anomalies occur only after such grounded truths have been taken into account. At the center of the problem of solvability are sinkless graphs: every solution for any graph G consists of the uniquely induced (grounded) $\alpha_{\overline{\varnothing}}$ composed with a solution for the ungrounded, sinkless residuum G° (recall that G° is the subgraph induced by $G^{\circ} = G \setminus (\overline{\varnothing} \cup E^{\smile}(\overline{\varnothing}))$).

Fact 4.1 *For any* G*:*

1. *$sinks(\mathsf{G}^{\circ}) = \varnothing$, and*

2. *$sol(\mathsf{G}) = \{\alpha \cup \alpha_{\overline{\varnothing}} \mid \alpha \in sol(\mathsf{G}^{\circ})\}$, hence also: $sol(\mathsf{G}) \neq \varnothing \Leftrightarrow sol(\mathsf{G}^{\circ}) \neq \varnothing$.*

So, although empirical contingency may influence the (non)paradoxical character of the actual discourse, eventually, it is always the ungrounded, non-empirical residuum of the discourse which determines such a character. In case of a "fully grounded" discourse, a dag with no infinite paths, G° is empty and the induced $\alpha_{\overline{\varnothing}}$ is the unique solution. But G° may also be empty when the graph contains cycles. In the example a) below, all statements obtain induced values as indicated; b) is paradoxical, as inducing leaves the unresolved liar, while c) is not paradoxical, having a truth-teller as the ungrounded residuum.

$$\text{a) This sentence is false and the Earth isn't round.}$$
$$\circlearrowright \mathbf{0} \longrightarrow \mathbf{1}$$

$$\text{b) This sentence is false and the Earth is round.}$$
$$\circlearrowright b \longrightarrow \mathbf{0} \longrightarrow \mathbf{1} \tag{4.2}$$

$$\text{c) This sentence is true and the Earth is round.}$$
$$\overline{c} \rightleftarrows c \longrightarrow \mathbf{0} \longrightarrow \mathbf{1}$$

Groundedness is sometimes taken to provide *the* source of definite and unavoidable truth-values. However, we have already seen several examples of discourses where statements that are not grounded still have truth-values that can be, intuitively, ascertained. A further example may be the statement x claiming both falsity and truth of the truth-teller z:

$$\mathsf{T} \qquad\qquad\qquad \mathcal{G}(\mathsf{T})$$

$$x \leftrightarrow \neg z \wedge \neg y$$
$$y \leftrightarrow \neg z$$
$$z \leftrightarrow \neg \overline{z}$$
$$\overline{z} \leftrightarrow \neg z$$
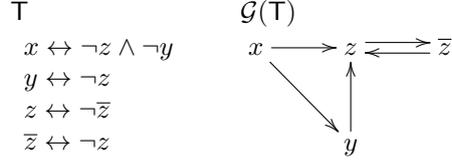
Accepting propositional logic only, such a statement is false, even if ungrounded. Indeed, T is consistent and proves classically $\neg x$. In our case, unlike in Kripke's least fixed-point, this conclusion is obtained by noting that no kernel of the solvable $\mathcal{G}(\mathsf{T})$ contains x, or by proving in PDL the possibility of $x = \mathbf{0}$, $\vdash \langle \neg x : \mathcal{G}(\mathsf{T}) \rangle$, and the impossibility of $x = \mathbf{1}$, $\vdash \neg \langle x : \mathcal{G}(\mathsf{T}) \rangle$.

Shortcomings of the least fixed-point approach have been addressed by proposing various other fixed-points as alternatives or as additions. But like every proliferation disease, this too poses the question where to stop. Possibilities suggested by Kripke were revised by the revision theory which, along with its own notion of stability, introduced a new plethora of different notions of validity and truth. This may possess some merit, and the comparison of different solutions is worthwhile, provided that it leads to a more definitive understanding and more definite theory of the phenomenon under question. In our case such problems do not arise – a discourse is paradoxical iff it corresponds to an inconsistent theory iff its graph has no kernel. All these questions are decided by PDL for the finitary discourses, and settled unambiguously by the general semantic results for the infinitary ones.

4.2.1 Non-empirical inducing

Groundedness is related to empirical contingency. As for the contingent liar a in (4.2), groundedness allows to dissolve the paradox, since Earth is round. But contingency need not be empirical. Consider F from Example 2.13.(2) of the liar f' contingent on the truth-teller e. If one stipulates that e is false then this discourse is paradoxical, if e is true it is not. This warrants the conclusion that not only the truth-teller e *could* be true but that it *must* be true in order for the discourse as a whole to function properly – T has a unique model, $e = \mathbf{1}$. This is a general phenomenon, illustrating again the holistic character of discourses: contingent paradoxes can give rise to definite truth-values of the ungrounded statements on which they depend. In light of the plausible meta-assumption that discourses which need not be paradoxical should *not* be designated as such, the presence of a potential paradox becomes *a fact* which, like empirical facts captured by $\overline{\varnothing}$, may force the truth-values of other statements.

There is, to our knowledge, no formal account of the semantical paradoxes that provides for this kind of reasoning. Note that although it may be referred to the holistic meta-imperative of avoiding paradoxes, it is simply a sound classical inference, concluding e (and $\neg f'$) from $f' \leftrightarrow \neg f' \wedge \neg e$. If the goal is to avoid paradox whenever possible then such inferences are appropriate. The example also seems to provide a counterargument against viewing the truth-teller itself

as pathological because its truth-value is arbitrary, depending only on itself. Apparently – and classically! – the truth-value of a truth-teller may depend on other statements not because it refers to them, but because they refer to the truth-teller.

Since this claim invokes a meta-principle (of global consistency), it creates a tension with the possibility of viewing the truth-teller $d - e$ as an isolated and independent subdiscourse, which has two local kernels. Limiting one's view to one's own business, without any consideration for larger issues, is a psychological possibility and it is not for logic to exclude it. The role of logic may be to describe such phenomena accurately and derive their unavoidable consequences. The provability $\vdash \langle \neg e : \mathsf{F} \rangle$ in PDL demonstrates only the possible falsity of e. But also necessary consequences of this option are provable, for instance, that it entails paradoxicality of $f' : \vdash \neg\langle \neg e \wedge \neg(\neg f' \wedge f') : \mathsf{F}\rangle$.

4.2.2   No discourse of truth-tellers is paradoxical

More can be said about the communities of truth-tellers, irrespectively of their (un)groundedness. Buridan's early solution to the liar and other paradoxes claimed that every statement, saying whatever it might be saying, says also "... and I am true." In terms of a graph, such a community X of truth-tellers involves, in addition to the actual edges between the nodes X, their copies $\overline{X} = \{\overline{x} \mid x \in X\}$, with a two cycle $x \leftrightarrows \overline{x}$ for every $x \in X$. Every kernel of the original graph, determines also a kernel of the new one (inducing $\overline{x} = \neg x$). But one obtains also the kernel $\overline{X}$, which means that all $x \in X$ can be $\mathbf{0}$, irrespectively of any connections between them. This makes any discourse almost void since, no matter the values of various (sub)statements, every statement can always be $\mathbf{0}$. No matter what various truth-tellers say about each other, they never become paradoxical, but they never become tautological either. This applies equally to any single truth-teller involved in any discourse: it is never paradoxical, because it can always be false. An argument asserting it's own truthfulness does not add any weight but, inadvertently, gives others the possibility to consider it false, irrespectively of any other circumstances.

Truth-teller appears also in the statement "This sentence is a paradox", with paradox understood dialetheically, i.e., x : "This sentence is true and false". The graph contains the truth-teller x claiming also its own falsity: $\overline{x} \rightleftarrows x \ $. Its only solution makes x false. (The statement is false also when the paradox is taken as a gap, neither true nor false, though this is no longer a truth-tellers community. Its graph becomes then $\overline{x} \gtrless x \overleftrightarrow{\ } x_2 \rightarrow x_1$ and has the only solution making x false.)

4.2.3   Every discourse of liars is paradoxical.

Unlike a truth-teller graph, with a 2-cycle at each vertex, a reflexive graph has no kernel since a node with a loop, $x \in E(x)$, can not belong to any kernel. In such a liar community everybody claims, in addition to whatever he may be claiming about others, also its own falsity. Every such a liar community

is paradoxical, for instance, each of the following discourses, where X is an arbitrary set with $|X| > 1$ and every $x \in X$ says:

(i) "I am lying." – this is just a collection of unrelated liars,

(ii) "Everybody, including me, is lying." or

(iii) "Everybody else is speaking truth but I am lying." or

(iv) "Every person with my eye-colour is lying." or

(v) "My right neighbour and both his neighbours are lying" (standing in a ring or an infinite line).

4.2.4 Accusations breed guilt

A graph G is weakly complete when its underlying, undirected graph $\underline{\mathsf{G}}$ is complete, i.e., when for each pair of distinct nodes $x \neq y$ in G, either $x \in E(y)$ or $y \in E(x)$. As is easy to see, *any kernel of a weakly complete graph – if it exists – is a single node x satisfying the kernel equation (2.7): $E^{\smile}(x) = G \setminus \{x\}$, [5].*

For instance, any company in which, for every two persons at least one accuses the other of lying, is a paradox, unless there is a person accused of lying by everybody except himself. Exactly one such person is telling the truth.

As a special case, for any set X with $|X| > 1$, let every $x \in X$ accuse everybody else (except himself) of lying. This gives a strongly (and hence also weakly) complete graph without loops, where each node satisfies the equation (2.7), and hence gives a possible kernel. Exactly one x is speaking the truth but it can be chosen arbitrarily, as can be the value for the truth-teller (which is the special case with $|X| = 2$.)

Also Yablo's graph, with the natural numbers $\mathbb{N}$ as nodes and the edge relation $E(x) = \{y \in \mathbb{N} \mid y > x\}$, is weakly complete. Its unsolvability follows, since for every $x \in \mathbb{N} : E^{\smile}(x) = \{y \in \mathbb{N} \mid y < x\} \neq \mathbb{N} \setminus \{x\}$.

The same argument shows paradox in any generalization of Yablo where, instead of $\mathbb{N}$, one takes integers, rationals, reals, or any other total order without greatest element.

5 Argumentation theory

Before concluding the paper, it seems appropriate to make a closer comparison with argumentation theory. Although its motivations and goals are different from ours, it moves within exactly the same formalism of directed graphs, exploring also (local) kernels and often similar intuitions to those underlying investigation of paradoxes.

At first, let us only remind that argumentation theory is a very wide landscape. Already in the seminal paper [18] several semantics were proposed and since then their number has only increased, along with the alternative interpretations and addressed questions. Our work concerns only the small corner of the argumentation landscape, using kernels and local kernels. These have been studied also in graph-theory and our work builds explicitly on this relation,

which seems to have been ignored by argumentation theory and has not been used earlier for the study of paradoxes.

Thus, for instance, Richardson's theorem can be now seen as expressing formally the intuition that vicious circularity is in fact *necessary* for (a finite semantic) paradox. Graph normal form and associated (local) kernels allowed us to distinguish vicious circularity from the innocent one, and even to determine classically possible truth-values of statements in an ungrounded discourse, to the point of inducing their values not from empirical evidence, but from the requirement of the consistency of the totality of discourse, cf. Section 4.2.1. Such inducing is seen as objectionable by various authors in argumentation theory, since it leads to accepting some arguments which have no intrinsic justification. PDL, being based on local kernels, allows us to explore some consequences of such a view, without abandoning the classical intuition that paradox arises precisely when discourse malfunctions. The insight we provide, in particular, is that a discourse can, under specific circumstances, malfunction even if there is no statement such that semantic failure is intrinsic to it; paradox arises from interaction, not from particular statements that can be identified as problematic. In the context of argumentation, many researchers seem to think that such a holistic diagnosis is unacceptable, that if a problem cannot be pinned down to individual arguments, then this is not a fault of argumentation structure, but rather of the semantics used to evaluate it.

This sentiment may motivate inventions of new semantics. It is clear, for instance, in the development of argumentation-semantics that seek to treat even cycles as being on par with odd ones. Arbitrariness ensues in both cases, the story goes, and so the semantics should not distinguish between them, see e.g., [1]. However, as we have seen, the distinction between odd and even cycles is abolutely *crucial* for the diagnosis of problematic cases. Conflating them, then, seems to amount to removing from semantic consideration what is the most important issue, namely the search for a better understanding of how semantic problems arise from interactions between arguments along odd cycles. Putting it more provocatively, conflating odd and even interactions asks us to overlook the difference between what is arbitrary and what is impossible. To us, coming at this from the point of view of paradoxes, this is an unavoidable – in fact, *the central* – distinction, both with regard to truth and to argumentation. This difference of opinions can be seen as reflecting the difference of motivations. An exchange of arguments may happen in an arbitrary way, and one would always like to decide which ones are to be preferred. A discourse, on the other hand, may be easier expected to conform to some normative rules, like that of overall coherence. The difference of origin leads here to the difference in the expectations of what the formal model should provide. Still, our claim would be that some arguments may by unsolvable, and it is worthwhile to specify which ones and why.

Even if we disagree with equating odd and even cycles, we should still mention that it has given rise to a technical concept that is quite elegant and has received a fair bit of attention from the argumentation community, namely, that of a stepwise inducing of accepted values along the dag of strongly connected

29

components of the argumentation digraph, the SCC-recursiveness, [2]. This notion refers to a framework for specifying argumentation semantics, where one starts from considering the dag of strongly connected components in the graph and chooses, in an arbitrary, way what to do on such components, starting from the terminal ones. Having assigned some semantic status to arguments in terminal components, inducing takes place along the acyclic part of the graph as prescribed by the classical, truth-functional rules (essentially, as given in Definition 2.9). This process is then iterated until every argument recieves some semantic status, we refer to [2] for technical details. The upshot is that by parameterizing this construction by a function that chooses arguments from strongly connected components, we arrive at a whole class of argumentation semantics that are, one might say, *intrinsically classical*, conforming to classical inducing whenever this provides a definite answer, but making some other, possibly non-classical, choice, otherwise.

The notion encompasses kernels and local kernels as well as several argumentation-semantics proposed earlier. It even allows the construction of semantics with completely *arbitrary* choices about which arguments to accept from strongly connected components. Being merely a tool for constructing actual semantics, it does not provide any specific proposal and is beyond comparison with any such. The most studied non-classical heuristic gives rise to $\mathcal{CF}2$ semantics, [1], which arises from allowing, for a strongly connected component, any maximal independent set of arguments, irrespectively of this set's ability to defend itself against attacks. Equivalently, in the context of discourses, the semantics allows us to regard some statements as false, even if what they say is considered true. This seems patently wrong from the point of view of propositional discourse, and also dubious from the point of view of argumentation. Still, we admit that given a starting point where ungroundedness is seen as always leading to arbitrariness, it does at least become a possible point of view. But even in this case, it is hardly an intrinsically justified solution, and the conceptual framework underlying it, straying from the classical intuitions underlying the original semantics introduced by Dung, opens instead the door for a great proliferation of non-classical suggestions. Notice also, for the case of $\mathcal{CF}2$, that as every graph has a maximal independent set, not only does classical logic dissapear as soon as we encounter an argument that fails to have an intrinsic status, but so does the notion of paradox dissapear from the semantics altogheter. Hence, the relevance for classical logic and paradox, addressed by us, of $\mathcal{CF}2$ or any other non-classical alternatives arising from SCC-recursiveness, is uncertain.[20]

A completely different relation to the work done on argumentation frameworks, concerns the modal character of our PDL, which is embedded in the existential quantification over local kernels in its definition of truth, 3.1. In-

---

[20]One can certainly envision more detailed distinctions within such a non-classical semantics, separating paradoxical and non-paradoxical discourses, but such distinctions remain to be proposed and investigated. Indeed, given that argumentation semantics try to determine successful arguments in *every possible* context, they will tend to disregard any distinction between proper and malfunctioning discourses, founding the concept of paradox.

deed, one can view PDL as a dialect of S5, since all assignments, determined
by local kernels, are "equally accessible" from each other. However, viewing
modality of PDL in this way, cripples it instead of clarifying. What is essential
for the involved notion of possibility are the available assignments. These are
determined by the (graphical) structure of the discourse. The fact that they all
are mutually and "equally accessible" is of marginal importance. This, so to say
"material" or "extensional" character of the modality, excluding also any nest-
ing of modalities (in a different way but with similar effects as in S5), seem to
make it a close relative of the informal, minimal modality of natural discourse.

This modal element, as the concern with mere consistency and not validity,
may capture most informal intuitions but is not what logicians understand by
modality. Since modal logic with Kripke semantics can be seen as a logic of
graphs (or of movements along directed edges of graphs), our use of graphs
might suggest turning to some existing modal logic. However, since each among
typical modal logics corresponds only to *some* subclass of graphs, a new variant
would be needed, allowing to model arbitrary referential structures. Such a
modal foundation has been proposed for argumentation networks, for instance,
in [19]. The basic form of the modal formula for a graph constructed there,
$\mu(\mathsf{G})$, bears close similarities to our graph normal form, GNF(G). Its advantage
arises in argumentation theory, since it allows to characterize logically various
kinds of extensions.[21] The expressive power, however, comes at a price. First,
in spite of the same basic form, $\mu(\mathsf{G})$ is significantly more complex than GNF(G).
More importantly, the semantic view involves at least three values and a level
of detail that eventually leads to very fine-grained distinctions, for instance,
between the following two discourses:

If the goal is to capture logically the exact structure of a graph, this is certainly
an advantage. But along with each discourse, its graph is given, so there seems
to be little need to duplicate its representation. Logic serves rather to obtain
some more abstract view of it. One can, for instance, view logic as a way of
capturing the elements of the discourse which are relevant for the truth-values
of its statements. In such a context, distinguishing between a discourse A, where
all statements are paradoxical, and B, where all statements are paradoxical, may
be more than what is actually called for.

Another approach is developed by Davide Grossi in [26, 25], where universal
modal logic is employed to give a logical characterization of argumentation se-
mantics. A formula $Adm(\phi)$, for instance, is constructed saying that the set of
arguments corresponding to ϕ is admissible (local kernel). This approach views
the directed graph as a modal frame, and it allows to import techniques from
modal logic to study various semantic notions from argumentation theory. It

---

[21]It can also be seen as a successful continuation of the attempts to determine where to
place the third value in pointer semantics for circular discourses, arising from [21].

does not, however, provide a logic based on any such semantics in particular. Possibly p, in the sense of there being an admissible set containing p, is not expressible except indirectly, by saying that there is some formula ϕ, corresponding to some set of arguments L, such that L is admissible and $p \in L$.

Theorems 3.8 and 3.10, on the other hand, establish a tight relation between the local kernel semantics of graphs, which underlies also $\mu(\mathsf{G})$, and the well-known logic Ł3. The applicability of Ł3 to the analysis of local kernels provides a novel perspective on this logic. (As a curiosity, let us note how Łukasiewicz's third semantic value returns thus, through graphs, to the modal applications, which motivated its introduction.) More importantly, the orderliness of PDL's sequent calculus allows to draw some conclusions about the structure of paradox from the properties of proofs of paradoxicality, as we saw in Section 3.3. Similar results are hard to expect for the mentioned modal works on argumentation theory, since an elegant and informative proof theory for many standard modal logics is missing (like, e.g., for the modal logic from [19], which extends K4 with Löb's axiom and more).

6   Concluding remarks

The holistic character of the discourse is reflected in that kernels can not be obtained by any straightforward, compositional rules.[22] Consistency is not a compositional property. At the same time, local kernels represent subdiscourses where (elements of) compositionality can be regained and from which meaningful information can be extracted even when other parts, or the totality of the discourse, are inconsistent. The logic PDL is paraconsistent in the sense that it handles meaningfully inconsistent discourses without any deductive explosion. But, at the same time, it is essentially classical, using only boolean truth-values and classical evaluation of connectives, with paradox (or lack of truth-value) appearing only as a non-functional consequence of the inconsistency of the discourse.

PDL can be seen as a formalization and completion of the project of logic of statements from [34], which asked exactly for such a propositional logic of discourses, represented as graphs in the same way as is done here.[23] Having now observed also the equivalence between a series of different problems, which so far have been considered in isolation, PDL can serve for addressing instances of each such separate problem: consistency of discourses, existence of kernels in digraphs, presence of semantic paradoxes, coherence of argumentation networks and even propositional satisfiability. (Applications to non-monotonic reasoning or logic programming were not considered here, but they are possible, too, as shown initially in [18], and later, for instance, in [12, 13, 14].) On the other

---

[22]SCC-recursiveness must not be misunderstood for such rules. It allows for an iterative construction (here, of kernels) only *between* strongly connected components, but tells one nothing how to obtain kernels of any such component.

[23]The only exception is the lack of the explicit truth-operator in PDL. Possibility of including such an operator remains to be investigated.

32

hand, the equivalence of different problems helps understanding each single one, in the light of its relations to others. In this respect, kernel theory has proved particularly enlightening, clarifying and making precise many intuitions about the nature of circularity.

On a more philosophical note, we saw that in practice there is no necessary opposition between the correspondence and coherence view of truth. The two can function in unison. External facts, sinks, induce values to some statements but, typically, do not cover the whole discourse. Once such inducing has taken place, there remains the problem of potential paradoxicality of the remaining part. For this ungrounded residuum, no sufficient, external criteria of truth are available and there remain only necessary, negative criteria demanding exclusion of undesirable effects, primarily, of inconsistency. Paradoxes appear only in this inner circle. Fact 4.1 captures precisely the intuition that while empirical contingency may contribute to dissolving them, it never creates any new paradoxes – in a finitary discourse, a paradox arises only due to some self-negation, a vicious circle. Furthermore, the problem of distinguishing between evil and innocent circularity has been extensively addressed in kernel theory. Applicability of its results to the anomalies of natural discourse seems a valuable insight, providing both precise results and an enhanced general understanding of the phenomenon.

Finally, defining PDL by the essentially classical conditions (2.5), we have avoided the problematic issue of the behaviour of logical connectives in the presence of paradox. Intuitively, saying that the liar is false, seems false, just as saying that it is true, seems false (unless one turns dialetheic). Yet from (3.15) we see that every paradox is a statement negating a paradox – the obvious example being the liar. So, sometimes, negation of a paradox is a paradox and other times, it is false? It might be possible to *impose* such a distinction between the *statements* of a discourse, resulting in a new, non-classical semantics, as the one given and elaborated by Gaifman in [21, 23, 22]. Its intuitive appeal seems to rest in large part on viewing paradox as a property of individual statements. If one insists on this, then it is only reasonable to let ⊥ result in a functional, compositional propagation of semantic values. However, the long lasting difficulties with agreeing on a single, stable set of rules determining such a functional propagation, let alone appearance, of paradox, should be very discouraging for this approach. The difficulties and proliferation of alternatives become more understandable, if paradox does not arise from any property of individual statements, but is a holistic effect of the totality of the discourse. Then it seems more appropriate to start with a logic that treats it as such and does not try to pin it down to particular statements. Also, although positive description of a paradox might be easy to ask for, it has proven notoriously difficult to obtain. It then seems more prudent to designate paradox negatively, as the limit (of consistency), found when we reach the point where classical intuitions, despite our best efforts, fail to provide any definitive answers.

Appendix: Proofs

Proposition 3.7 *Given a graph* G, *we have:*
a) If $L \in Lk(\mathsf{G})$ *then* $\alpha_{\overline{L}} \models_L \mathcal{D}(\mathsf{G})$ *and;*
b) If $\alpha \models_L \mathcal{D}(\mathsf{G})$ *for* $\alpha : G \to \{\mathbf{1}, \mathbf{0}, \bot\}$ *then* $\overline{\varnothing} \subseteq \alpha^{\mathbf{1}} \in Lk(\mathsf{G})$

PROOF. a) Assume $L \in Lk(\mathsf{G})$ and consider arbitrary $x \leftrightarrow \bigwedge_{y \in E(x)} \neg y \in \mathcal{D}(\mathsf{G})$. Assume towards contradiction that $\alpha_{\overline{L}} \not\models_L x \leftrightarrow \bigwedge_{y \in E(x)} \neg y$. We let $\overline{\alpha}_{\overline{L}}$ denote the evaluation of complex formulae obtained from $\alpha_{\overline{L}}$ according to tables (3.6). Then we have $\alpha_{\overline{L}}(x) \neq \overline{\alpha}_{\overline{L}}(\bigwedge_{y \in E(x)} \neg y)$. If $\alpha_{\overline{L}}(x) = \mathbf{1}$ this inequality means that there is one $y \in E(x)$ such that $\alpha_{\overline{L}}(y) \in \{\mathbf{1}, \bot\}$, impossible by the fact that $\overline{L}$ is a local kernel (which requires, for all $y \in E(x)$, $y \in E^{\smile}(\overline{L})$, i.e. $\alpha_{\overline{L}}(y) = \mathbf{0}$). If $\alpha_{\overline{L}}(x) = \mathbf{0}$ we must then have, for every $y \in E(x)$, $\alpha_{\overline{L}}(y) \in \{\mathbf{0}, \bot\}$ but this is also ruled out by the fact that $\overline{L}$ is a local kernel (which requires existence of some $y \in E(x)$ such that $\alpha_{\overline{L}}(y) = \mathbf{1}$). The last possibility is that $\alpha_{\overline{L}}(x) = \bot$ in which case there are two possibilities. 1) We have some $y \in E(x)$ such that $\alpha_{\overline{L}}(y) = \mathbf{1}$. This contradicts $x \notin E^{\smile}(L)$ (required since we have $\alpha_{\overline{L}}(x) = \bot$). 2) For all $y \in E(x)$ we have $\alpha_{\overline{L}}(y) = \mathbf{0}$. This means $x \in sinks(\mathsf{G} \setminus (\overline{L} \cup E^{\smile}(\overline{L}))$, impossible by Definition 2.9 of $\overline{L}$.
b) Assume $\alpha \models_L \mathcal{D}(\mathsf{G})$. We show that $\alpha^{\mathbf{1}}$ is a local kernel. We show first that $\alpha^{\mathbf{1}}$ is independent. Assume towards contradiction that it is not. Then there are $x, y \in \alpha^{\mathbf{1}}$ with $y \in E(x)$. So we have $\alpha(x) = \alpha(y) = \mathbf{1}$ and from inspecting the tables (3.6) we see that $\overline{\alpha}(\bigwedge_{y \in E(x)} \neg y) = \mathbf{0}$. In particular, we have $\alpha(x) \neq \overline{\alpha}(\bigwedge_{y \in E(x)} \neg y)$, contrary to hypothesis. Assume towards contradiction that $\alpha^{\mathbf{1}}$ is not locally absorbing. Then there is some $x \in \alpha^{\mathbf{1}}$ with $y \in E(x)$ such that $E(y) \cap \alpha^{\mathbf{1}} = \varnothing$. Since $\alpha(z) \neq \mathbf{1}$ for all $z \in E(x)$, by $\alpha^{\mathbf{1}}$ being independent, this means that $\alpha(y) = \bot$ and that $\overline{\alpha}(\bigwedge_{y \in E(x)} \neg y) = \bot \neq \alpha(x) = \mathbf{1}$, contrary to hypothesis. To show that $\overline{\varnothing} \subseteq \alpha^{\mathbf{1}}$ is a simple proof by induction over definition 2.9. For the basis, if $x \in sinks(\mathsf{G})$, i.e. $x \in \varnothing_1$, then $x \leftrightarrow \mathbf{1} \in \mathcal{D}(\mathsf{G})$. Then is is clear that $x \in \alpha^{\mathbf{1}}$. The inductive step is also trivial. □

Theorem 3.8 $\models \langle \Gamma : \mathsf{G} \rangle$ *iff there is some* $\alpha : G \to \{\mathbf{1}, \mathbf{0}, \bot\}$ *such that* $\alpha \models_L \mathcal{D}(\mathsf{G})$ *and* $\alpha \models_L \Gamma$.

PROOF. $\Rightarrow$) Assume that $\models \langle \Gamma : \mathsf{G} \rangle$. Then by Definition 3.1 there is some local kernel $L \in Lk(\mathsf{G})$ such that $L \models \langle \Gamma : \mathsf{G} \rangle$. We have from Proposition 3.7.a) that $\alpha_{\overline{L}} \models_L \mathcal{D}(\mathsf{G})$. We show $\alpha_{\overline{L}} \models_L \Gamma$ by induction on the complexity of Γ. We take its complexity to be the sum of the complexity of its formulae divided by $|\Gamma|$. The basis is for Γ a collection of literals. Then $\Gamma^+ \subseteq L$ and $\Gamma^- \subseteq E^{\smile}(L)$, so for all $x \in \Gamma^+$ we have $\alpha_{\overline{L}}(x) = \mathbf{1}$ and for all $y \in \Gamma^-$ we have $\alpha_{\overline{L}}(y) = \mathbf{0}$ (remember that $L \subseteq \overline{L}$). It follows from inspecting tables (3.6) that $\alpha_{\overline{L}} \models \Gamma$. The inductive steps are easy. For instance, if $\models \langle \Gamma : \mathsf{G} \rangle$ and there is $\neg\neg A \in \Gamma$ that has maximal complexity among formulae of Γ, then we form Γ' which is like Γ except that A replaces $\neg\neg A$. Γ' has lower complexity than Γ and, obviously from Definition 3.1, $\models \langle \Gamma' : \mathsf{G} \rangle$. So by IH we get $\models_L \Gamma'$. Consulting tables (3.6) we see that this gives us $\models_L \Gamma$ so we are done. The cases for $\neg(A \wedge B)$ and $A \wedge B$ are equally

easy.

$\Leftarrow$) Assume $\alpha \models_L \mathcal{D}(\mathsf{G})$ and $\alpha \models_L \Gamma$. We have $\alpha^{\mathbf{1}} \in Lk(\mathsf{G})$ from 3.7.b) and obtain $\alpha^{\mathbf{1}} \models \langle \Gamma : \mathsf{G} \rangle$ by induction on the complexity of Γ, measured as in the proof of $\Rightarrow$). From this $\models \langle \Gamma : \mathsf{G} \rangle$ follows by Definition 3.1. The basis is for Γ a collection of literals. Consulting tables (3.6), we see that for all $x \in \Gamma^+$, we have $\alpha(x) = \mathbf{1}$ so $x \in \alpha^{\mathbf{1}}$. For all $y \in \Gamma^-$, on the other hand, we have $\alpha(y) = \mathbf{0}$. Since $\alpha \models_L \mathcal{D}(\mathsf{G})$, we have $\alpha \models_L y \leftrightarrow \bigwedge_{z \in E(y)} \neg z$, meaning $\alpha(y) = \overline{\alpha}(\bigwedge_{z \in E(y)} \neg z)$ (where $\overline{\alpha}$ is the evaluation of α according to tables 3.6). From tables (3.6) we see that there must then be some $z \in E(y)$ such that $\alpha(z) = \mathbf{1}$, meaning $y \in E^{\smile}(\alpha^{\mathbf{1}})$. It follows from Definition 3.1 that $\alpha^{\mathbf{1}} \models \langle \Gamma : \mathsf{G} \rangle$. The inductive steps are straightforward. For instance, if there is some $A \wedge B \in \Gamma$ that has maximal complexity among formulae of Γ, we form Γ' which is like Γ except that we replace $A \wedge B$ by A and B. Then $\models_L \Gamma'$ and Γ' has smaller complexity than Γ so by IH $\alpha^{\mathbf{1}} \models \langle \Gamma' : \mathsf{G} \rangle$. It follows immediately from Definition 3.1 that $\alpha^{\mathbf{1}} \models \langle \Gamma : \mathsf{G} \rangle$ and we are done. The cases of $\neg\neg A$ and $\neg(A \wedge B)$ are equally easy. $\qquad\square$

Soundness and completeness of PDL

Soundness and completeness follow easily from the following simple lemma giving us the compositionality we need with respect to admissibility in graphs.

Lemma 6.1 *For any graph G and $a \in G$ we have:*

    (1) $\models \langle \Gamma, a : \mathsf{G} \rangle$ *iff* $\models \langle \Gamma, \{\neg b \mid b \in E(a)\} : \mathsf{G} \setminus out(a) \rangle$

    (2) $\models \langle \Gamma, \neg a : \mathsf{G} \rangle$ *iff for some* $b \in E(a)$, $\models \langle \Gamma, b : \mathsf{G} \rangle$

PROOF. (1) $\Rightarrow$) Assume Γ, a is admissible in G and let $L \subseteq G$ be a local kernel witnessing to Γ and containing a. Clearly, L is a local kernel also in $\mathsf{G} \setminus out(a)$. Now, since $a \in L$ it follows that $E(a) \subseteq E^{\smile}(L)$, so $\Gamma \cup \{\neg b \mid b \in E(a)\}$ is indeed admissible (in both G and $\mathsf{G} \setminus out(a)$)

$\Leftarrow$) Assume $\Gamma \cup \{\neg b \mid b \in E(a)\}$ is admissible in $\mathsf{G} \setminus out(a)$ and let $L \subseteq G$ be an arbitrary local kernel in $\mathsf{G} \setminus out(a)$ witnessing to this fact. Then for every $b \in E(a)$ we have $E(b) \cap L \neq \varnothing$ so $L \cup \{a\}$ is a local kernel in G (as well as in $\mathsf{G} \setminus out(a)$)

(2) $\Rightarrow$) Let $L \subseteq G$ be a local kernel witnessing to the admissibility of $\Gamma, \neg a$ in G. Then, for some $b \in E(a)$, we have $b \in L$. So Γ, b is admissible in G.

$\Leftarrow$) Assume that there is some $b \in E(a)$ such that Γ, b is admissible. Let $L \subseteq G$ be a witness. Then L also witness to the admissibility of $\Gamma, \neg a$ in G. $\qquad\square$

This lemma establishes soundness and invertibility of the only rules from PDL that are not essentially classical. The rest is easily verified, yielding

Theorem 6.2 PDL *is sound and all its rules are invertible.*

PROOF. The standard sequent rules for the composite formulae in $\Theta \vdash \Phi$ are trivially invertible, as are the rules for non-atomic basic $\langle \Gamma : \mathsf{G} \rangle$ (which form

a one-sided sequent system for propositional logic). Lemma 6.1 established soundness and invertibility of the four rules for literals in Γ. We only have to show that the two axiom schemata are valid:

(1) $\Theta, \langle \Gamma, \neg a : \mathsf{G} \rangle \vdash \Phi$ for some $a \in sinks(\mathsf{G})$.

To show $\Theta, \langle \Gamma, \neg a : \mathsf{G} \rangle \models \Phi$, it suffices to show that $\not\models \langle \neg a : \mathsf{G} \rangle$, by Definition 3.1. By Definition 3.1, this amounts to the nonexistence of a local kernel L of G containing a successor of a. But since a is a sink in G, no such L exists.

(2) $\Theta \vdash \langle \Gamma : \mathsf{G} \rangle, \Phi$ for some $\Gamma \subseteq sinks(\mathsf{G})$.

To show $\Theta \models \langle \Gamma : \mathsf{G} \rangle, \Phi$, it suffices to show $\models \langle \Gamma : \mathsf{G} \rangle$. Since Γ is a collection of atomic expressions this amounts to showing that there is a local kernel L in G such that $\Gamma \subseteq L$. But $sinks(\mathsf{G})$ is such a local kernel in G so the claim follows.
$\square$

Completeness of PDL follows now by the standard line of reasoning, demonstrating invalidity of any unprovable sequent. We say that a sequent $\Theta \vdash \Phi$ is *reduced* when Θ and Φ contain only atomic formulae, i.e., every $\langle \Gamma : \mathsf{G} \rangle \in \Theta \cup \Phi$ contains only literals and, moreover, literals over sinks of G, i.e.,

$$\Gamma = \{a \mid a \in \Gamma^+ \subseteq sinks(\mathsf{G})\} \cup \{\neg b \mid b \in \Gamma^- \subseteq sinks(\mathsf{G})\}.$$

We first argue that, for any sequent, the rules suffice to create a proof-tree with all leafs reduced.

Trivially, the top level rules and rules for composite Γ suffice to create a proof-tree where all leafs have the form $\Theta \vdash \Phi$ with Θ and Φ being collections of atomic expressions $\langle \Gamma : \mathsf{G} \rangle$, i.e., each Γ being a collection of literals. Now, we employ the rules for literals, as long as there is some $a \in \Gamma$ or $\neg a \in \Gamma$ with $E(a) \neq \varnothing$, i.e. as long as the sequent is not reduced. For any finite graph G, it is clear that by employing these rules we will eventually reach a stage where all sequents have been reduced. If (i)

$a \in \Gamma$ is not a sink, an application of the rule $(\vdash a)$, resp., $(a \vdash)$, makes it a sink. If (ii) $\neg a \in \Gamma$ is not a sink, then an application of the rule $(\vdash \neg)$, resp., $(\neg \vdash)$, replaces it by all its out-neighbours with positive polarity, for which case (i) applies in the next round.

Theorem 3.14 *System* PDL *is sound and complete:* $\Theta \vdash \Phi$ *iff* $\Theta \models \Phi$.

PROOF. We show that reduced, non-axiomatic $\Theta \vdash \Phi$, is invalid. We have:

(1) $\forall \langle \Gamma_T : \mathsf{G}_T \rangle \in \Theta : \Gamma_T^- = \varnothing$ and (2) $\forall \langle \Gamma_F : \mathsf{G}_F \rangle \in \Phi : \Gamma_F^- \neq \varnothing$.

(1) follows since $\Theta \vdash \Phi$ is reduced, so for all $\langle \Gamma_T : \mathsf{G}_T \rangle \in \Theta : \Gamma_T^+ \cup \Gamma_T^- \subseteq sinks(\mathsf{G}_T)$. Since the sequent is not axiomatic, we must have $\Gamma_T^- = \varnothing$. Consequently, $\Gamma_T \subseteq sinks(\mathsf{G}_T)$ and, since $sinks(\mathsf{G}_T) \in Lk(\mathsf{G}_T)$, so $\models \langle \Gamma_T : \mathsf{G}_T \rangle$.

(2) follows since, as before, $\Gamma_F^+ \cup \Gamma_F^- \subseteq sinks(\mathsf{G}_F)$ and, since the sequent is not axiomatic, $\Gamma_F \not\subseteq sinks(\mathsf{G}_F)$. Consequently, $\Gamma_F^- \neq \varnothing$ (since atoms from this set are negated in Γ_F). So there is some $a \in \Gamma_F^- \subseteq sinks(\mathsf{G}_F)$, i.e. $\neg a \in \Gamma_F$ while $a \in sinks(\mathsf{G}_F)$. It follows that $\not\models \langle \Gamma_F : \mathsf{G}_F \rangle$.

Having obtained $\models \langle \Gamma_T : \mathsf{G}_T \rangle$ for all $\langle \Gamma_T : \mathsf{G}_T \rangle \in \Theta$ and $\not\models \langle \Gamma_F : \mathsf{G}_F \rangle$ for all $\langle \Gamma_F : \mathsf{G}_F \rangle \in \Phi$, we conclude by Definition 3.1 that $\Theta \not\models \Phi$. Invertibility of all the rules ensures that if such a reduced sequent is obtained as a leaf in a proof

tree from some initial sequent S, then also S is invalid. Invertibility was shown in Theorem 6.2 and here we also established soundness of the system. $\qquad\square$

References

[1] Pietro Baroni and Massimiliano Giacomin. Solving semantic problems with odd-length cycles in argumentation. In ThomasDyhre Nielsen and Nevin-Lianwen Zhang, editors, *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, volume 2711 of *Lecture Notes in Computer Science*, pages 440–451. Springer Berlin Heidelberg, 2003.

[2] Pietro Baroni, Massimiliano Giacomin, and Giovanni Guida. Scc-recursiveness: a general schema for argumentation semantics. *Artif. Intell.*, 168(1):162–210, 2005.

[3] Jon Barwise and Lawrence Moss. *Vicious Circles: On the Mathematics of Non-Wellfounded Phenomena*. CSLI, Stanford, 1996.

[4] J. C. Beall. Is Yablo's paradox non-circular? *Analysis*, 63(1):176–187, 2001.

[5] Marc Bezem, Clemens Grabmayer, and Michał Walicki. Expressive power of digraph solvability. *Annals of Pure and Applied Logic*, 2011. [to appear].

[6] Jean-Yves Béziau. A sequent calculus for Łukasiewicz's three-valued logic based on suszko's bivalent semantics. *Bulletin of the Section of Logic*, 28(2):89–97, 1998.

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

[8] Vašek Chvátal. On the computational complexity of finding a kernel. Technical Report CRM-300, Centre de Recherches Mathématiques, Univeristé de Montréal, 1973. `http://users.encs.concordia.ca/~chvatal`.

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

[10] Roy Cook. There are non-circular paradoxes (but Yablo's isn't one of them). *The Monist*, 89:118–149, 2006.

[11] Nadia Creignou. The class of problems that are linearly equivalent to satisfiability or a uniform method for proving np-completeness. *Theoretical Computer Science*, 145:111–145, 1995.

[12] Yannis Dimopoulos and Vangelis Magirou. A graph theoretic approach to default logic. *Information and Computation*, 112:239–256, 1994.

[13] Yannis Dimopoulos, Vangelis Magirou, and Christos H. Papadimitriou. On kernels, defaults and even graphs. *Annals of Mathematics and Artificial Intelligence*, 20:1–12, 1997.

[14] Yannis Dimopoulos and Alberto Torres. Graph theoretical structures in logic programs and default theories. *Theoretical Computer Science*, 170(1-2):209–244, 1996.

[15] Sylvie Doutre. *Autour de la sématique préférée des systèmes d'argumentation*. PhD thesis, Université Paul Sabatier, Toulouse, 2002.

[16] Pierre Duchet. Graphes noyau-parfaits, II. *Annals of Discrete Mathematics*, 9:93–101, 1980.

[17] Pierre Duchet and Henry Meyniel. Une généralisation du théorème de Richardson sur l'existence de noyaux dans les graphes orientés. *Discrete Mathematics*, 43(1):21–27, 1983.

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

[19] Dov Gabbay. Modal provability foundations for argumentation networks. *Studia Logica*, 93(2-3):181–198, 2009.

[20] Dov Gabbay and Leendert van der Torre, editors. *New Ideas in Argumentation Theory: Special Issue*, volume 93 (2-3). *Studia Logica*, 2009.

[21] Haim Gaifman. Operational pointer semantics: solution to self-referential puzzles. In Moshe Vardi, editor, *Theoretical Aspects of Reasoning about Knowledge*, pages 43–59. Morgan Kauffman, 1988.

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

[23] Haim Gaifman. Pointers to propositions. In Andre Chapuis and Anil Gupta, editors, *Circularity, Definition and Truth*, pages 79–121. Indian Council of Philosophical Research, 2000.

[24] Hortensia Galeana-Sánchez and Victor Neumann-Lara. On kernels and semikernels of digraphs. *Discrete Mathematics*, 48(1):67–76, 1984.

[25] Davide Grossi. Argumentation in the view of modal logic. In Peter McBurney, Iyad Rahwan, and Simon Parsons, editors, *ArgMAS*, volume 6614 of *Lecture Notes in Computer Science*, pages 190–208. Springer, 2010.

[26] Davide Grossi. On the logic of argumentation theory. In Wiebe van der Hoek, Gal A. Kaminka, Yves Lespérance, Michael Luck, and Sandip Sen, editors, *AAMAS*, pages 409–416. IFAAMAS, 2010.

[27] Saul Kripke. Outline of a theory of truth. *The Journal of Philosophy*, 72(19):690–716, 1975.

[28] Victor Neumann-Lara. Seminúcleos de una digráfica. Technical report, Anales del Instituto de Matemáticas II, Universidad Nacional Autónoma México, 1971.

[29] H. Prakken and G. Vreeswijk. Logics for deafeasible argumentation. In Dov Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume 4, pages 219–318. Kluwer Academic Publishers, 2002.

[30] Graham Priest. Yablo's paradox. *Analysis*, 57:236–242, 1997.

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

[32] Roy Sorensen. Yablo's paradox and kindred infinite liars. *Mind*, 107:137–155, 1998.

[33] John von Neumann and Oscar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944 (1947).

[34] Michał Walicki. Reference, paradoxes and truth. *Synthese*, 171:195–226, 2009.

[35] Michał Walicki and Sjur Dyrkolbotn. Finding kernels or solving SAT. 2010. http://www.ii.uib.no/~michal/KernelSAT.pdf.

[36] Stephen Yablo. Paradox without self-reference. *Analysis*, 53(4):251–252, 1993.