The Review of Symbolic Logic
Volume 18, Number 1, March 2025

REFERENCE DIGRAPHS OF NON-SELF-REFERENTIAL PARADOXES

MING HSIUNG
School of Philosophy and Social Development
South China Normal University

Abstract. All the known non-self-referential paradoxes share a reference pattern of Yablo's paradox in that they all necessarily contain infinitely many sentences, each of which refers to infinitely many sentences. This raises a question: Does the reference pattern of Yablo's paradox underlie all non-self-referential paradoxes, just as the reference pattern of the liar paradox underlies all finite paradoxes? In this regard, Rabern et al. [J Philos Logic 42(5): 727–765, 2013] prove that every dangerous acyclic digraph contains infinitely many points with an infinite out-degree. Building upon their work, this paper extends Rabern et al.'s result to the first-order arithmetic language with a primitive truth predicate, proving that all reference digraphs for non-self-referential paradoxes contain infinitely many sentences of infinite out-degree (called "social sentences"). We then strengthen this result in two respects. First, among these social sentences, infinitely many appear in one ray. Second, among these social sentences, infinitely many have infinitely many out-neighbors, none of which will eventually get to a sink. These observations provide helpful information towards the following conjecture proposed by Beringer and Schindler [Bull. of Symb. Logic 23(4): 442–492, 2017]: every dangerous acyclic digraph contains the Yablo digraph as a finitary minor.

§1. Introduction. It has long been realized that there is a close relationship between the truth-theoretic paradoxes and reference patterns of sentences. Through a series of graph-theoretic analyses of paradoxes, people recently have gained a fairly clear understanding of the above relationship. First, it has been proved that if a finite set of sentences is paradoxical, there must be a circular reference pattern between these sentences. In terms of (di)graph theory, a finite reference digraph (for a set of sentences) is dangerous (in the sense that these sentences lead to a paradox) iff it contains a directed cycle. From this, we can prove that a finite reference digraph is dangerous iff it contains a subdivision of the liar digraph as a sub-digraph.[1] Thus, the characterization problem of reference patterns has been completely solved for finite paradoxes.

As for the infinite paradoxes, the characterization problem is far more challenging. What complicates the matter is that for some infinite sets of paradoxical sentences, there may not be any cyclic pattern between these sentences at all. That is, there are (infinite) paradoxes that are non-self-referential or non-circular. The first and also the

Received: July 3, 2023.
2020 *Mathematics Subject Classification*: Primary 03A05, Secondary 05C20.
Key words and Phrases: reference digraph, self-reference, social point, truth predicate, Yablo's paradox.

[1] This result is first proved by Rabern et al. [22, p. 751] in a specific infinitary propositional language. Rabern et al.'s discovery is extended to a first-order language by Beringer & Schindler [2, p. 474].

© The Author(s), 2024. Published by Cambridge University Press on behalf of The Association for Symbolic Logic.

doi:10.1017/S1755020324000200

https://doi.org/10.1017/S1755020324000200 Published online by Cambridge University Press

best-known example is Yablo's paradox.[2] After the invention of Yablo's paradox, more and more non-self-referential paradoxes have been constructed.[3] Moreover, people have been able to prove formally that Yablo's paradox and other related ones are indeed non-self-referential by the aforementioned graph-theoretic methods.[4]

With the advancement of research, many people have noticed that all the known non-self-referential paradoxes share a reference pattern of Yablo's paradox in that they all necessarily contain infinitely many sentences, each of which refers to infinitely many sentences. So, people [2, p. 474] ask: is this just an accidental phenomenon, or does the reference pattern of Yablo's paradox underlie all non-self-referential paradoxes, just as the reference pattern of the liar paradox underlies all finite paradoxes?

Concerning the above question, we must mention a theorem proved by Rabern et al. [22, pp. 756–757]: every dangerous acyclic digraph contains infinitely many points with an infinite out-degree. As far as the author knows, this is the first significant result related to the above question. However, Rabern et al.'s result, as Beringer & Schindler [2, p. 445] point out, is formulated in an infinitary propositional language so that the definition of the so-called reference graphs, an highly crucial part of Rabern et al.'s [22] framework, "no longer yields satisfactory results when we move to first-order languages." (ibid., p. 445).

This paper extends Rabern et al.'s result to the standard language of studying truth and paradoxes, namely, the first-order arithmetic language with a primitive truth predicate. Within this language, we employ the notion of dependency relation introduced by Leitgeb [19] to define reference digraphs for paradoxes. Based on this framework, we prove that every non-self-referential paradox contains infinitely many sentences (called "social sentences"), each of which depends on infinitely many sentences. This result is an equivalent formulation of Rabern et al.'s result in the first-order language. We then strengthen this result by proving that there are infinitely many of these social sentences in a ray. Furthermore, as we will see, the social sentences appearing in a ray can meet an even stronger condition: each of them depends on infinitely many sentences, none of which eventually gets to a sink.

Our study yields some insights into the reference patterns of non-self-referential paradoxes, revealing a structural similarity shared by all such paradoxes with Yablo's paradox. Our observation prompts a conjecture proposed by Beringer & Schindler: "A reference [di]graph is dangerous iff it contains a subdivision of the liar-graph as a subgraph or the Yablo-graph as a final minor." The non-trivial aspect of this

---

[2] On the one hand, the non-self-referentiality of Yablo's paradox is considered so self-evident by many (including Yablo himself) that it requires no more explanation. For instance, in Yablo [29, 30], although Yablo takes what he constructs as a "paradox without self-reference", he never explains a single word about its non-self-referentiality. On the other hand, Priest [21], among others, disagrees that Yablo's paradox is non-self-referential. Priest's point of view is irrelevant to our current study. The circularity of Yablo's paradox he discusses is not associated with the referential structure of sentences on which the present article focuses. We refer the reader to Leitgeb [18] for more information.

[3] Among all the known constructions, we highlight Cook's [5, p. 770] unwinding, a method that can systematically generate paradoxes similar to Yablo's paradox. A specific unwinding variant, now known as the $\exists\forall$-unwinding variant, is given by Yablo [31, p. 144]. This method has been further developed by Cook [6] himself, Schlenker [23], and Hsiung [15] in various directions. Additionally, it is worth noting that Butler [4] provides a procedure to construct continuum-many variants of Yablo's paradox.

[4] See, for instance, [5, p. 770], [19, p. 170], and [14, pp. 902–903].

https://doi.org/10.1017/S1755020324000200 Published online by Cambridge University Press

conjecture lies in the claim that all dangerous acyclic digraphs contain the Yablo digraph as a finitary minor. While this paper has not yet proven this conjecture, our research endeavors to contribute helpful information that may ultimately lead to its confirmation.

As mentioned above, we carry out our research in the first-order language of Peano arithmetic with a primitive truth predicate. In this language, we define the notion of reference digraphs by employing the notion of semantic dependence relation proposed by Leitgeb [19] (§2). We establish a version of Rabern et al.'s theorem in the first-order arithmetic language and give a strengthening involved in rays in §3. Then, in §4, we prove that those infinitely many social sentences can meet the stronger condition mentioned above. Based on the above results, we give a series of characterizations about the dangerous digraphs in §5. Finally, in the concluding, we briefly discuss a potential approach to Beringer & Schindler's conjecture based on the existing findings.

Graph-theoretic Preliminaries. Let $\mathcal{G}$ be a digraph (without parallel edges), that is, a pair $\langle D, \prec\rangle$, where is D is a non-empty set and $\prec$ is a binary relation on D. For two points u and v of D, $u \prec v$ is read as "u 'see' v". Let $u_0\, u_1\, \ldots\, u_l$ be a finite sequence of points of $\mathcal{G}$. If for all $0 \leq i < l$, at least one of u_i and u_{i+1} can see the other, then we call this sequence a *walk* in $\mathcal{G}$. u_0 and u_l are also called the two *endpoints* of this walk. l is called its *length*. This walk is *directed*, if u_i sees u_{i+1} for all $0 \leq i < l$. In that case, u_l is *reachable* from u_0. A walk is a *path*, if no point of it is repeated except possibly its two endpoints. Note that a (directed) walk always contains a (directed) path connecting two endpoints of this walk.

A walk is *closed* if its endpoints are the same point. A *directed cycle* is a closed directed path. A *loop* is a directed cycle of length 1. A digraph is *acyclic*, if it contains no directed cycles (equivalently, no closed directed walks). An acyclic digraph, that is, a directed acyclic graph, is also called a DAG. A digraph is *loop-free*, if it contains no loops.

Whenever u sees v, we also say v is an *out-neighbor* of u (and u is an *in-neighbor* of v). The *out-degree* of a point is the size (cardinality) of the set of its out-neighbors. A point of out-degree zero is called a *sink*. A digraph is *locally finite*, if each of its points has a finite out-degree.

A digraph $\mathcal{G} = \langle D, \prec\rangle$ is *conversely well-founded*, if every non-empty subset S of D has a $\prec$-maximal element, that is, an element of S having no out-neighbor in S. A digraph is *conversely ill-founded*, if it is not conversely well-founded. A *ray* $\mathcal{G}$ is an infinite $\prec$-increasing sequence of points in $\mathcal{G}$, $u_0 \prec u_1 \prec \ldots$. Note that under the axiom of choice, a digraph $\mathcal{G}$ is conversely ill-founded, iff there is a ray in $\mathcal{G}$.

The relation $\prec^*$ is the *transitive closure* of $\prec$, if it meets the requirements: $u \prec^* v$, iff v is reachable from u by a directed path. For any u, we use $\prec^*(u)$ for the set $\{v \mid u \prec^* v\}$. Also, we use $\preccurlyeq$ for the reflexive closure of $\prec$. That is, $u \preccurlyeq v$, iff $u \prec v$ or $u = v$. Thus, $\preccurlyeq^*$ is the transitive closure of $\preccurlyeq$. It can be seen as the reflexive and transitive closure of $\prec$.

Note that whenever D' is a subset of D, the sub-digraph $\langle D', \prec \restriction_{D'}\rangle$ is usually denoted by $\langle D', \prec\rangle$ for brevity's sake.

§2. Paradoxes and Their Reference Digraphs. Let $\mathcal{L}_T$ be the first-order language of Peano arithmetic (PA) with a unary predicate symbol T. It is well known that we can formulate the paradoxical and other pathological sentences in $\mathcal{L}_T$ by a routine

https://doi.org/10.1017/S1755020324000200 Published online by Cambridge University Press

diagonal method.[5] For instance, we can construct a sentence L such that $L \leftrightarrow \neg T \ulcorner L \urcorner$ is provable in PA. L is the very liar sentence. $T \ulcorner L \urcorner$ is a shorthand for $T(\ulcorner L \urcorner)$, in which $\ulcorner L \urcorner$ is the numeral corresponding to the Gödel number of L. If no confusion arises, we also use it to denote the Gödel number of L itself. Another important example is the set of sentences $Y_0, Y_1, \ldots$ satisfying that for all number n, $Y_n \leftrightarrow \forall x\,(x > \bar{n} \rightarrow \neg T \ulcorner Y_{\dot{x}} \urcorner)$ is provable in PA.[6] $Y_0, Y_1, \ldots$ are Yablo's sentences.

In this article, we always fix the standard structure $\mathfrak{N}$ of natural numbers as the ground model for $\mathcal{L}_T$. So, by a structure for $\mathcal{L}_T$, we mean a pair $\langle \mathfrak{N}, X \rangle$, where $X \subseteq \mathbb{N}$ is an interpretation for T. We denote the valuation of a sentence A by $\mathcal{V}_X(A)$, which is a shorthand for $\mathcal{V}_{\langle \mathfrak{N}, X \rangle}$. When $\mathcal{V}_X(A) = 1/0$, we say A is true/false under the interpretation X. Sometimes, we identify a sentence with its Gödel number. For instance, for any set Σ of sentences, the "Σ" in the notation $\mathcal{V}_\Sigma$ is not Σ itself but the set of the Gödel numbers of sentences in Σ.

By Tarski's undefinability theorem of truth, it is impossible to find an interpretation $X \subseteq \mathbb{N}$ of T such that all instances of Tarski's T-schema $T \ulcorner A \urcorner \leftrightarrow A$ are true under X. The reason is that the liar sentence L or any other paradox Σ would lead to a contradiction provided that the instances of the T-schema $T \ulcorner A \urcorner \leftrightarrow A$ obtained from L or the sentences in Σ were true. So, the following is a folk definition of paradoxicality.[7]

Definition 2.1. *$X \subseteq \mathbb{N}$ is a* truth predicate *for a set Σ of sentences, if $\mathcal{V}_X(T \ulcorner A \urcorner) = \mathcal{V}_X(A)$ holds for any $A \in \Sigma$. Σ is* paradoxical, *if no $X \subseteq \mathbb{N}$ is a truth predicate for Σ.*

By definition, we can see that the liar sentence L (i.e., the singleton $\{L\}$) is paradoxical, so is the set of Yablo's sentences Y_n's. We leave the details to the reader.

Next, we introduce Leitgeb's [19] semantic dependence relation, which is a fundamental concept of studying the (self-)reference relation between sentences in this article. From now on, we use $\Sigma(A)$ instead of $\mathcal{V}_\Sigma(A)$ to make the notation more compact.

Definition 2.2 (Leitgeb [19]). *A sentence* A depends on *a set Σ of sentences, if for any sets Γ_1, Γ_2 of sentences, whenever $\Sigma \cap \Gamma_1 = \Sigma \cap \Gamma_2$, $\mathcal{V}_{\Gamma_1}(A) = \mathcal{V}_{\Gamma_2}(A)$.*

Intuitively, A depends on Σ, iff the truth value of A is only relevant to "the presence or absence of the sentences that are contained in Φ [here, Σ] in/from the extension of the truth predicate." [19, p. 160] The following are three primary properties about the dependence relation (ibid., p. 161).

Lemma 2.3. *The dependence relation has the following properties*:

(1) *Any sentence depends on $\mathcal{L}_T$, i.e., the set of all the sentences.*
(2) *If A depends on Σ and $\Sigma \subseteq \Sigma'$, then A depends on Σ'.*
(3) *If A depends on both Σ and Σ', then A depends on $\Sigma \cap \Sigma'$.*

For example, we can verify that the liar sentence depends on $\{L\}$, which is the smallest one among all the sets on which L depends. In this sense, L is *essentially*

[5] See, for instance, [3, pp. 53–54] for more information about this method.
[6] $\bar{n}$ is the numeral denoting the the number n. $\dot{x}$ is Feferman's [9] dot notation, by which we allow the scope of the quantifier $\forall x$ covers the formula Y_x, even though Y_x hides behind a closed term.
[7] See also [14, p. 893].

https://doi.org/10.1017/S1755020324000200 Published online by Cambridge University Press

depends on $\{L\}$ (ibid., p. 162). Also, for any number n, Y_n essentially depends on the set $\{Y_k \mid k > n\}$.

DEFINITION 2.4 (Beringer & Schindler [1]). *f is a* dependence function *on $\mathcal{L}_T$, if it is a function which assigns to any sentence a set of sentences, such that A depends on $f(A)$ for any A. $\prec_f$, a binary relation on $\mathcal{L}_T$, is defined by: $A \prec_f B$, iff $B \in f(A)$. The digraph $\langle \mathcal{L}_T, \prec_f \rangle$ is called the* reference digraph (of $\mathcal{L}_T$) induced from f.[8]

Let Σ be a set of sentences. The digraph $\langle \Sigma, \prec_f \rangle$, or more precisely, $\langle \Sigma, \prec_f \restriction_\Sigma \rangle$, is called a *reference digraph* of Σ (induced from f). What is more, $\prec_f$ is called a *reference relation* on Σ. The following theorem establishes a fundamental connection between a paradox and its reference digraphs.[9]

THEOREM 2.5. *If a set of sentences is paradoxical, then any of its reference digraphs is conversely ill-founded.*

Proof. Let Σ be a set of sentences. Suppose that there is a dependence function f such that $\langle \Sigma, \prec_f \rangle$ is conversely well-founded. We prove that Σ is not paradoxical. Note that since the relation $\prec_f$ is conversely well-founded on Σ, we have a rank function on Σ, say ρ, along its converse.

For any ordinal α, define inductively Σ_α as follows.[10] First, let Σ_0 be Σ. Second, let $\Sigma_{\alpha+1}$ be the set of sentences A in Σ such that $\Sigma_\alpha(A) = 1$. Finally, for a limit α, let Σ_α be the limit inferior of the sequences $\langle \Sigma_\beta \mid \beta < \alpha \rangle$, that is,

$$\Sigma_\alpha = \{A \mid \exists \beta < \alpha \forall \gamma (\beta \leq \gamma < \alpha \rightarrow A \in \Sigma_\gamma)\}.$$

CLAIM. *For any $A \in \Sigma$, whenever $\alpha > \rho(A)$, $A \in \Sigma_\alpha$, iff $A \in \Sigma_{\rho(A)+1}$.* □

If this claim is proved, then let θ be the least upper bound of the ordinals $\rho(A)$ for all $A \in \Sigma$. Then, for all $A \in \Sigma$, $A \in \Sigma_{\theta+1}$, iff $A \in \Sigma_{\theta+2}$. Hence, $\Sigma_{\theta+1}(T \ulcorner A \urcorner) = \Sigma_{\theta+1}(A)$ holds for all $A \in \Sigma$. Consequently, Σ is not paradoxical.

Proof of Claim. The proof is a transfinite induction on the rank of A. □

First, when $\rho(A) = 0$, we can see $f(A) \cap \Sigma = \varnothing$. Also, for any $\alpha \geq 0$, Σ_α is a subset of Σ, and so, $f(A) \cap \Sigma_\alpha = \varnothing$. By the definition of dependence relation, we have $\Sigma_\alpha(A) = \Sigma(A)$. Hence, for all $\alpha \geq 1$, $A \in \Sigma_\alpha$, iff $A \in \Sigma_1$.

Next, fix a sentence A in Σ, and suppose that the claim is true for all $B \in \Sigma$ with $\rho(B) < \rho(A)$. We prove the claim for A. We consider two cases. If $\alpha = \gamma + 1$, it suffices to prove that for any $\gamma \geq \rho(A)$, $\Sigma_\gamma(A) = \Sigma_{\rho(A)}(A)$. For this, we only need to prove that $f(A) \cap \Sigma_\gamma = f(A) \cap \Sigma_{\rho(A)}$. Fix $B \in f(A)$ arbitrarily. Since $A \prec_f B$, $\rho(B) < \rho(A)$. Also, we know $\gamma \geq \rho(A)$, and so $\gamma > \rho(B)$. Now, by the induction hypothesis, $B \in \Sigma_\gamma$, iff $B \in \Sigma_{\rho(B)+1}$. By the induction hypothesis again, the latter is equivalent to $B \in \Sigma_{\rho(A)}$. We thus obtain $B \in \Sigma_\gamma$, iff $B \in \Sigma_{\rho(A)}$. The desired equation follows.

---

[8] The notion of reference digraph is initially introduced by Rabern et al. [22, p. 737] in the context of an infinite propositional language. Our definition is a slightly modified version of Beringer & Schindler [1]. See also [2, p. 453].

[9] An equivalent result with a different form has been found by Rabern et al. [22, pp. 749–750] in a setting of infinitary propositional language. Note that their proof is non-constructive due to Zorn's lemma.

[10] This definition applies a revision rule proposed independently by Gupta [10, p. 10] and Herzberger [13, p. 68].

https://doi.org/10.1017/S1755020324000200 Published online by Cambridge University Press

If α is a limit greater than $\rho(A)$, then by definition of Σ_α, it suffices to find an ordinal $\beta < \alpha$ such that for all γ with $\beta \leq \gamma < \alpha$, $A \in \Sigma_\gamma$ holds, iff $A \in \Sigma_{\rho(A)+1}$. Let β be the supremum of all the ordinal $\rho(B)+1$ for $B \in f(A)$, that is, $\beta = \bigcup_{B \in f(A)}(\rho(B)+1)$. Note that $A \prec_f B$, and so $\rho(B) < \rho(A)$, and $\rho(B)+1 \leq \rho(A)$. It follows $\beta \leq \rho(A) < \alpha$. Now, if $\beta \leq \gamma < \alpha$, then for all $B \in f(A)$, $\gamma > \rho(B)$. By the inductive hypothesis, $B \in \Sigma_\gamma$, iff $B \in \Sigma_{\rho(B)+1}$. On the other hand, we also have $B \in \Sigma_{\rho(A)}$, iff $B \in \Sigma_{\rho(B)+1}$ because $\rho(A) > \rho(B)$. To sum up, we find an ordinal $\beta < \alpha$, such that for any $B \in f(A)$ and any γ with $\beta \leq \gamma < \alpha$, $B \in \Sigma_\gamma$, iff $B \in \Sigma_{\rho(A)}$. Thus, $\Sigma_\gamma \cap f(A) = \Sigma_{\rho(A)} \cap f(A)$, which implies $\Sigma_\gamma(A) = \Sigma_{\rho(A)}(A)$. Consequently, β satisfies that for all γ with $\beta \leq \gamma < \alpha$, $A \in \Sigma_\gamma$, iff $A \in \Sigma_{\rho(A)+1}$.

We now relate the reference digraph to the notion of self-reference. The following definition of self-reference, given by Hsiung [14, p. 895], is a generalization of Leitgeb's [19, p. 168] notion of (direct) self-reference.

DEFINITION 2.6. *A set* Σ *of sentences is* self-referential, *if each reference digraph of* Σ *contains at least a* closed directed walk. *That is, for any dependence function* f, *there are sentences* $A_1, \ldots A_n$ *in* Σ, *such that* $A_1 \prec_f A_2 \prec_f \ldots \prec_f A_n$ *and* $A_1 = A_n$.

It can be proved that Σ is self-referential, iff each of its reference digraphs contains at least a *closed directed path* (a *directed cycle*), that is, a closed directed walk in which none of the points is repeated except the two endpoints ($A_1 = A_n$). See [14, pp. 895–896] for the proof.

Suppose $\prec$ be a conversely ill-founded relation on a set Σ, then by the axiom of choice, it is not hard to find an infinite increasing sequence of elements of Σ, say, $A_0 \prec A_1 \prec A_2 \prec \ldots$. If Σ is finite, this sequence must contain a directed cycle. A *ray* is an infinite increasing sequence in which no points are repeated. The following corollary is straightforward by Theorem 2.5.

COROLLARY 2.7.

(1) *If a paradoxical set of sentences is finite, it is self-referential.*
(2) *If a paradoxical set of sentences is non-self-referential, each of its reference digraphs contains a ray.*

In the proof of Theorem 2.5, $\Sigma_{\theta+1}$, the set witnessing the non-paradoxicality of Σ, is a subset of Σ. So, as to Corollary 2.7 (1), if a finite set Σ is non-self-referential, we can find a *subset* of Σ such that it is a truth predicate for Σ. This result has been proved in [14, Theorem 1, pp. 896–897]. Theorem 2.5 is indeed a generalization of the latter. We will further strengthen Corollary 2.7 (2) by showing that for a non-self-referential paradox, each of its referential digraphs contains a ray in which there are infinitely many "social" sentences (to be made clear later).

We close this section with three examples. First, the liar sentence is evidently self-referential. Second, the set of Yablo's sentences Y_n ($n \geq 0$) is non-self-referential. For this, only note that for any numbers n, m and any dependence function f, $Y_n \prec_f Y_m$, iff $n < m$. Thus, it is impossible to find a finite sequence $Y_{n_1} \prec_f Y_{n_2} \prec_f \ldots \prec_f Y_{n_k}$ with $Y_{n_1} = Y_{n_k}$. At last, we consider sentences M_n ($n \geq 0$) such that $M_0 \leftrightarrow \exists x \neg T \ulcorner M_{\dot{x}} \urcorner$ and $M_{n+1} \leftrightarrow T \ulcorner M_n \urcorner$ ($n \geq 0$) are provable in PA. Let us call these sentences "McGee's sentences" for the construction of them is due to McGee [20, p. 400]. M_0 essentially depends on $\{M_n \mid n \geq 0\}$, and so the set of McGee's sentences is self-referential. Also, it is paradoxical. We leave the details to the reader.

https://doi.org/10.1017/S1755020324000200 Published online by Cambridge University Press

§3. Social Sentences. The three examples given at the end of the previous section represent three categories of paradoxes. The liar sentence is a typical finite paradox (i.e., a paradoxical set containing only finitely many sentences). By the first result of Corollary 2.7, all finite paradoxes are self-referential. Both Yablo's paradox and McGee's paradox are infinite paradoxes. Also, we notice a difference between them: the former contains infinitely many sentences which only depend on infinite sets, while there is only such sentence, namely M_0, in the latter. All known non-self-referential paradoxes, such as various variants of Yablo's paradox, satisfy the property we just point out about Yablo's paradox. So, a natural question is whether there is any infinite paradox in which only finitely many sentences depend on infinite sets.

DEFINITION 3.1. *Let $\mathcal{G} = \langle D, \prec \rangle$ be a digraph. A point $u \in D$ is* social *in $\mathcal{G}$, if* u *is a point of infinite out-degree in $\mathcal{G}$. That is to say,* u *has infinitely many out-neighbors in D. $\mathcal{G}$ is* locally finite, *if no point of D is social in $\mathcal{G}$.*

The following theorem gives a negative answer to the above question.

THEOREM 3.2. *If a paradoxical set is non-self-referential, any of its reference digraphs contains infinitely many social sentences.*

Let us say that a set of sentences is locally finite, if it has a locally finite reference digraph. One specific case of Theorem 3.2 is as follows. Our proof of Theorem 3.2 depends on the proof of this specific case.

THEOREM 3.3 (Hsiung [14]). *If a paradoxical set of sentences is locally finite, it is self-referential.*

Note that Theorem 3.2 is a strengthening of Theorem 3.3, as the latter is equivalently to say that for any non-self-referential paradoxical set of sentences, its reference digraphs always contain at least one social sentence. Later, we will further strengthen Theorem 3.2 to assert that for any non-self-referential paradoxical set of sentences, its reference digraphs always contain infinitely many social sentences, each of which can "see" infinitely many ungrounded points (see Theorem 4.3).

The proof of Theorem 3.3 is based on the first result of Corollary 2.7. We refer the reader to [14, pp. 897–898] for details.

Before proving Theorem 3.2, we still need to utilize two lemmata.

LEMMA 3.4. *Suppose $\langle \Sigma, \prec \rangle$ is a reference digraph of Σ and Δ is a $\prec$-closed subset of Σ. If Δ has a truth predicate, then it has one which is a subset of Δ.*

Proof. Let X be a truth predicate for Δ, we prove that $X \cap \Delta$ is also a truth predicate for Δ. First note that when $A \in \Delta$, $A \in X \cap \Delta$, iff $A \in X$. The latter is equivalent to $\mathcal{V}_X(A) = 1$. Let $\prec = \prec_f$ for some dependence function f. Then, by $\prec$-closedness of Δ, $f(A)$ is included in Δ. Thus, A depends on Δ. We get $\mathcal{V}_X(A) = 1$, iff $\mathcal{V}_{X \cap \Delta}(A) = 1$. Consequently, $A \in X \cap \Delta$, iff $\mathcal{V}_{X \cap \Delta}(A) = 1$. □

The following lemma is a version of a result proved by Rabern et al. [22, p. 755] in a propositional logic setting. In our proof of the lemma, the idea of applying Zorn's lemma to closed sets is credited to Rabern et al., but in the setting of first-order language, we need to rely on the properties of dependence relations to find the required truth predicate.

https://doi.org/10.1017/S1755020324000200 Published online by Cambridge University Press

LEMMA 3.5. *Suppose $\langle \Sigma, \prec \rangle$ is a reference digraph of Σ. If Σ is paradoxical, then there is a non-empty subset Δ of Σ such that any non-empty $\prec \restriction_\Delta$-closed subset of Δ is also paradoxical.*

Proof. Let $\prec = \prec_f$ for some dependence function. Assume that for any non-empty subset Δ of Σ, there is a non-empty $\prec \restriction_\Delta$-closed subset of Δ such that it is not paradoxical. We consider the collection of all pairs $\langle \Theta, X \rangle$ such that Θ is a $\prec \restriction_\Delta$-closed subset of Δ for some $\Delta \subseteq \Sigma$ and X is a truth predicate for Θ. By Lemma 3.4, we can suppose $X \subseteq \Theta$. We denote this collection by $\mathcal{S}$. Define a binary relation $\leq$ on $\mathcal{S}$ as follows: $\langle \Theta, X \rangle \leq \langle \Theta', X' \rangle$, iff $\Theta \subseteq \Theta'$ and $X = X' \cap \Theta$. We can easily prove that $\leq$ is a partial order on $\mathcal{S}$. We leave the details to the reader.

To apply Zorn's lemma to $\langle \mathcal{S}, \leq \rangle$, we first note that $\mathcal{S}$ is non-empty since the pair $\langle \emptyset, \emptyset \rangle$ is apparently a member of it. Now, suppose $\langle \Theta_0, X_0 \rangle \leq \ldots \leq \langle \Theta_\alpha, X_\alpha \rangle \leq \ldots$ is a chain in $\langle \mathcal{S}, \leq \rangle$, where α belongs to an index set I. For each $\alpha \in I$, we have that Θ_α is $\prec \restriction_{\Delta_\alpha}$-closed, with Δ_α being a subset of Σ. We define Θ_I as the union of all Θ_α with $\alpha \in I$, and Δ_I as the union of all Δ_α with $\alpha \in I$. Moreover, we define a subset X_I of Θ_I as follows: for any $A \in \Theta_I$, let α be the smallest ordinal with $A \in \Theta_\alpha$, then include A in X_I if and only if $A \in X_\alpha$.

Now, we claim that $\langle \Theta_I, X_I \rangle$ is an upper bound for the given chain.

It is clear that Θ_I is $\prec \restriction_{\Delta_I}$-closed. $\langle \Theta_I, X_I \rangle$ is an upper bound of all $\langle \Theta_\alpha, X_\alpha \rangle$ for $\alpha \in I$. For this, note that by our definition of X_I, for any $A \in \Theta_\alpha$, $A \in X_I$, iff $A \in X_\alpha$. That is, $X_I \cap \Theta_\alpha = X_\alpha$.

We will next verify that X_I is a truth predicate for Θ_I. Let's fix $A \in \Theta_I$. As mentioned earlier, let α be the least ordinal with $A \in \Theta_\alpha$. Due to the closedness of $\prec \restriction_{\Delta_\alpha}$, we know that $f(A)$ is a subset of Θ_α, and so A depends on Θ_α. Therefore, we have the following chain of equivalences:

$$\begin{aligned} A \in X_I, \text{ iff } & A \in X_\alpha, \quad \text{(by definition of } \alpha) \\ & \text{iff } \mathcal{V}_{X_\alpha}(A) = 1, \quad (X_\alpha \text{ is a truth predicate for } \Theta_\alpha) \\ & \text{iff } \mathcal{V}_{X_I \cap \Theta_\alpha}(A) = 1, \quad (X_\alpha = X_I \cap \Theta_\alpha) \\ & \text{iff } \mathcal{V}_{X_I}(A) = 1. \quad (A \text{ depends on } \Theta_\alpha). \end{aligned}$$

As a result, every chain in $\langle \mathcal{S}, \leq \rangle$ has an upper bound. Zorn's lemma gives us a maximal element in $\langle \mathcal{S}, \leq \rangle$, denoted as $\langle \Theta, X \rangle$. We claim that $\Theta = \Sigma$. If this claim is true, our proof is complete.

Let's assume the contrary, i.e., $\Theta \neq \Sigma$. In that case, the set $\Sigma \setminus \Theta$ is non-empty. Consequently, we can find a non-empty subset Θ' of $\Sigma \setminus \Theta$ such that Θ' is closed under $\prec \restriction_{\Sigma \setminus \Theta}$ and is not paradoxical. Let $X' \subseteq \Theta'$ be a truth predicate for Θ'.

We prove that $\langle \Theta \cup \Theta', X \cup X' \rangle$ is a member of S. First, it is clear that $\Theta \cup \Theta'$ is clearly $\prec \restriction_\Delta$-closed. To see that $X \cup X'$ is a truth predicate for $\Theta \cup \Theta'$, we observe that if $A \in \Theta$, then A depends on Θ. Thus, since $(X \cup X') \cap \Theta = X$, we have $\mathcal{V}_{X \cup X'}(A) = \mathcal{V}_X(A)$. Similarly, if $A \in \Theta'$, we can get $\mathcal{V}_{X \cup X'}(A) = \mathcal{V}_{X'}(A)$. From this, it follows that for all $A \in \Theta \cup \Theta'$, we have $A \in X \cup X'$, iff $\mathcal{V}_{X \cup X'}(A) = 1$. Consequently, we can conclude that $X \cup X'$ is a truth predicate for $\Theta \cup \Theta'$.

Since $X' \cap \Theta = \emptyset$, we can immediately see that $\langle \Theta, X \rangle \leq \langle \Theta \cup \Theta', X \cup X' \rangle$ and they are not equal. $\langle \Theta, X \rangle$ is not maximal, a contradiction. □

Proof of Theorem 3.2. Suppose $\langle \Sigma, \prec \rangle$ is a reference digraph of Σ and it is a DAG. It suffices to prove that for any natural number n, if there are at most n social points in

https://doi.org/10.1017/S1755020324000200 Published online by Cambridge University Press

$\langle \Sigma, \prec \rangle$, then Σ is not paradoxical. We prove this result by induction. The case for $n = 0$ is Theorem 3.3.

Consider the case when there are at most $n + 1$ social points in $\langle \Sigma, \prec \rangle$. We will apply Lemma 3.5 to the proof. For this, for any non-empty subset Δ of Σ, fix a point in Δ, namely A. Consider the set $\prec^*(A)$. If it contains no social points, then by Theorem 3.3, this set is not paradoxical. Thus, $\prec^*(A) \cap \Delta$ is a non-paradoxical and $\prec_\Delta$-closed subset of Δ. If it contains at least a social point, fix one of these social points and let it be B. Now the set $\prec^*(B)$ contain at most n social point since B is out of this set. By the inductive hypothesis, $\prec^*(B)$ is not paradoxical. Again, we obtain a $\prec_\Delta$-closed subset of Δ, namely, $\prec^*(B) \cap \Delta$, which is non-paradoxical. To sum up, for any non-empty subset Δ of Σ, we can find a non-empty $\prec \restriction_\Delta$-closed subset of Δ such that it is not paradoxical. By Lemma 3.5, Σ is not paradoxical. □

From Theorem 3.2, together with Lemma 3.5, we can deduce the following important consequence.

THEOREM 3.6. *Suppose $\langle \Sigma, \prec \rangle$ is a reference digraph of Σ. If Σ is a non-self-referential and paradoxical set, then there is a ray $A_0 \prec^* A_1 \prec^* A_2 \prec^* \ldots$ in some sub-digraph $\langle \Delta, \prec \rangle$ of $\langle \Sigma, \prec \rangle$ such that for all natural number n, A_n is social in $\langle \Delta, \prec \rangle$.*

Proof. Recall that by $\langle \Delta, \prec \rangle$, we mean $\langle \Delta, \prec \restriction_\Delta \rangle$. By Lemma 3.5, we can fix a sub-digraph $\langle \Delta_0, \prec \rangle$ of $\langle \Sigma, \prec \rangle$ such that any non-empty $\prec \restriction_{\Delta_0}$-closed subset of Δ_0 is also paradoxical. Since $\langle \Delta_0, \prec \rangle$ is a DAG, by Theorem 3.2, we can find a social point in $\langle \Delta_0, \prec \rangle$. Let it be A_0 and Let Θ_0 denote $\preccurlyeq^*(A_0) \cap \Delta_0$. Note that since A_0 is social in $\langle \Delta_0, \prec \rangle$, A_0 is also social in $\langle \Theta_0, \prec \rangle$. Since Θ_0 is a $\prec \restriction_{\Delta_0}$-closed set in $\langle \Delta_0, \prec \rangle$, it must be paradoxical by our choice of the digraph $\langle \Delta_0, \prec \rangle$. Hence, by Lemma 3.5 again, we can fix a sub-digraph $\langle \Delta_1, \prec \rangle$ of $\langle \Delta_0, \prec \rangle$ such that any non-empty $\prec \restriction_{\Delta_1}$-closed subset of Δ_1 is also paradoxical. By Theorem 3.2 again, we can fix a social point in $\langle \Delta_1, \prec \rangle$, namely A_1. Let Θ_1 denote $\preccurlyeq^*(A_1) \cap \Delta_1$. Note that A_1 is social in $\langle \Theta_1, \prec \rangle$. Besides, since $\Theta_1 \subseteq \Theta_0$, A_1 is also social in $\langle \Theta_0, \prec \rangle$.

Repeating the above process, we can obtain a ray $A_0 \prec^* A_1 \prec^* A_2 \prec^* \ldots$, and a sequence $\Delta_0 \supsetneq \Delta_1 \supsetneq \Delta_2 \supsetneq \ldots$ such that for all $n \geq 0$, $\Theta_n = \preccurlyeq^*(A_n) \cap \Delta_n$ is $\prec \restriction_{\Delta_n}$-closed in Δ_n, and A_n is social in $\langle \Theta_n, \prec \rangle$. It is clear that every A_n is social in $\langle \Theta_0, \prec \rangle$ since $\Theta_0 \supseteq \Theta_n$. □

Theorem 3.2 tells us that there are infinitely many social sentences in any reference digraph for this paradox. Now by Theorem 3.6, we know, among the infinitely many social sentences we find in any reference digraph for this paradox, there are infinitely many appearing in a ray. So, Theorem 3.6 provides us with more information than Theorem 3.2. By the way, Theorem 3.6 is also a strengthening of Corollary 2.7 (2).

We conclude this section with a generalization of Theorem 2.5.

DEFINITION 3.7. *Let $\mathcal{G} = \langle D, \prec \rangle$ be a digraph. The* social part *of $\mathcal{G}$ is the set of all points $u \in D$ with $u \preccurlyeq^* v$ for some social point* v *in D.*

THEOREM 3.8. *If a paradoxical set is non-self-referential, then for any reference digraph, its social part is conversely ill-founded.*

Proof. Let Σ be a non-self-referential set of sentences. Then there exists a dependence function f such that the corresponding reference digraph $\langle \Sigma, \prec_f \rangle$ is acyclic. Suppose there is a reference digraph $\langle \Sigma, \prec_{f'} \rangle$ whose social part is conversely well-founded. Let f'' be a function on $\mathcal{L}_T$ defined by $f''(A) = f(A) \cap f'(A)$. Then f'' is a dependence

https://doi.org/10.1017/S1755020324000200 Published online by Cambridge University Press

function on $\mathcal{L}_T$. By replacing f with f'', we can suppose f meets the conditions (i) $\langle \Sigma, \prec_f \rangle$ is a DAG, and (ii) the social part of $\langle \Sigma, \prec_f \rangle$ are conversely well-founded.

We need to prove that Σ is not paradoxical. To this end, we must find a set, say Γ_θ, such that $\Gamma_\theta(T\ulcorner A\urcorner) = \Gamma_\theta(A)$ holds for all sentence A in Σ.

We now divide Σ into two parts. Let Σ_s be (the domain of) the social part of the digraph digraph $\langle \Sigma, \prec_f \rangle$. Let Σ_c be $\Sigma \setminus \Sigma_s$, i.e., the complement of Σ_s relative to Σ. Note that Σ_c is closed to $\prec_f$. That is, if $A \in \Sigma_c$ and $A \prec_f B$, then $B \in \Sigma_c$.

By our supposition (ii), the restriction of $\langle \Sigma, \prec_f \rangle$ to Σ_s, i.e., $\langle \Sigma_s, \prec_f \restriction_{\Sigma_s} \rangle$, is conversely well-founded. Hence, there is a rank function ρ over the set Σ_s for (the converse) of $\prec_f$.

Since every sentence in Σ_c is locally finite and non-self-referential, by Theorem 3.3, there exists Γ such that for all $A \in \Sigma_c$, $\Gamma(T\ulcorner A\urcorner) = \Gamma(A)$. For each ordinal α, we inductively define a set Γ_α as follows: $\Gamma_0 = \Gamma$, $\Gamma_{\alpha+1} = \{A \in \Sigma \mid \Gamma_\alpha(A) = 1\}$, and for a limit α, Γ_α is the limit inferior of the sequences $\langle \Gamma_\beta \mid \beta < \alpha \rangle$.

CLAIM 1. *For any $A \in \Sigma_c$ and any ordinal α, $A \in \Gamma_\alpha$, iff $A \in \Gamma$.* □

Proof of Claim 1. Fix $A \in \Sigma_c$, we prove by transfinite induction on α. The base case $\alpha = 0$ is evident. For the case $\alpha = \gamma + 1$, it is sufficient to show $\Gamma_\gamma(A) = \Gamma(A)$. For this, we only need to prove $\Gamma_\gamma \cap f(A) = \Gamma \cap f(A)$. Whenever $B \in f(A)$, we have $A \prec_f B$. Since Σ_c is closed to $\prec_f$, $A \in \Sigma_c$ implies $B \in \Sigma_c$. By the induction hypothesis, $B \in \Gamma_\gamma$, iff $B \in \Gamma$. The desired result follows immediately. □

If α is a limit, let $\beta = \bigcup_{B \in f(A)} (\rho(B) + 1)$. Then, for any $B \in f(A)$, if $\beta \leq \gamma < \alpha$, then $\rho(B) < \gamma < \alpha$, and by the induction hypothesis, $B \in \Gamma_\gamma$, iff $B \in \Gamma$. Therefore, $\Gamma_\gamma \cap f(A) = \Gamma \cap f(A)$. For any γ with $\beta \leq \gamma < \alpha$, we obtain $A \in \Gamma_{\gamma+1}$, iff $A \in \Gamma$. We can conclude that $A \in \Gamma_\alpha$, iff $A \in \Gamma$.

CLAIM 2. *For any $A \in \Sigma_s$ and any $\alpha > \rho(A)$, $A \in \Gamma_\alpha$, iff $A \in \Gamma_{\rho(A)+1}$.*

Proof of Claim 2. Again, the proof is a transfinite induction on the rank of $A \in \Sigma_s$. First, when $\rho(A) = 0$, we need to prove for any $\alpha > 0$, $A \in \Gamma_\alpha$, iff $A \in \Gamma_1$. In this case, A must be a social sentence, and $f(A) \subseteq \Sigma_c$. If $\alpha = \gamma + 1$, then for any $B \in f(A)$, by Claim 1, $B \in \Gamma_\gamma$, iff $B \in \Gamma$. It follows $\Gamma_\gamma \cap f(A) = \Gamma \cap f(A)$. Thus, $\Gamma_\gamma(A) = \Gamma(A)$, and so $A \in \Gamma_\alpha$, iff $A \in \Gamma_1$. If α is a limit, let $\beta = \bigcup_{B \in f(A)} (\rho(B) + 1)$. Then, by Claim 1 again, we can prove for any γ with $\beta \leq \gamma < \alpha$, $\Gamma_\gamma \cap f(A) = \Gamma \cap f(A)$, and thus, $\Gamma_\gamma(A) = \Gamma(A)$. We can obtain that $A \in \Gamma_\alpha$, iff $A \in \Gamma_1$.

Next, suppose the claim is true for any $B \in \Sigma_s$ with $\rho(B) < \rho(A)$. We must prove it is true for $A \in \Sigma_s$. The proof is still a transfinite induction on α, which is the same as the basis case $\rho(A) = 0$, except that we appeal to the inductive hypothesis instead of Claim 1. We omit the details.

Let θ be the least upper bound of the ordinals $\rho(A) + 1$ for all $A \in \Sigma_s$. By the above two claims, we can conclude that for any sentence A of Σ, if $\alpha \geq \theta$, then $A \in \Gamma_\alpha$, iff $A \in \Gamma_\theta$. In particular, $A \in \Gamma_{\theta+1}$, iff $A \in \Gamma_\theta$. That is, $\Gamma_\theta(T\ulcorner A\urcorner) = \Gamma_\theta(A)$. □

§4. Strongly Social Sentences. In this section, we continue to strengthen the theorems we have obtained in the previous section.

We first introduce the notion of grounded points.

https://doi.org/10.1017/S1755020324000200 Published online by Cambridge University Press

DEFINITION 4.1. *Let $\mathcal{G} = \langle D, \prec \rangle$ be a digraph. A point u of D is called a* grounded *point in $\mathcal{G}$, if $\preccurlyeq^*(u)$ is conversely well-founded*; *otherwise, it is* ungrounded. *The* grounded part *of $\mathcal{G}$ is the set of all grounded point in $\mathcal{G}$.*

Note that a *sink* is a point of out-degree zero. Thus, a sentence is grounded in a reference digraph, iff any of the walks starting from it eventually reaches a sink.[11] Recall that in a digraph $\mathcal{G} = \langle D, \prec \rangle$, a point $u \in D$ is a *social* point, if u can see infinitely many points.

DEFINITION 4.2. *Let $\mathcal{G} = \langle D, \prec \rangle$ be a digraph. $u \in D$ is a* strongly social *point in $\mathcal{G}$, if u can see infinite many ungrounded points in $\mathcal{G}$. A* weakly social *point (in $\mathcal{G}$) is a social point whose sociality (in $\mathcal{G}$) is not strong.*

We must emphasize that the notion of sociality given by Definition 4.2 is relative to a specific digraph. In particular, the distinction between strong social points and weakly social ones makes sense only if we have specified to which digraph these points belong. For some sentences, when we say that they are strongly (or weakly) social, we mean that they are so in some reference digraph fixed in the context.

The following is a strengthening of Theorem 3.2.

THEOREM 4.3. *If a paradoxical set is non-self-referential, any of its reference digraphs contains infinitely many strongly social sentences.*

The following theorem is a specific case of Theorem 4.3. We first prove it by König infinity lemma before proving Theorem 4.3. The proof is a generalization of the proof given by Hsiung [14, pp. 897–898].

THEOREM 4.4. *If a paradoxical set is non-self-referential, the social sentences in any of its reference digraphs cannot be all weak.*

Proof. Let Σ be a non-self-referential set of sentences. Suppose it has a reference digraph in which all social sentences are weak. So, we can find a dependence function f such that $\langle \Sigma, \prec_f \rangle$ whose social points are all weak. Since Σ is non-self-referential, we can, as we do in Theorem 3.2, further suppose $\langle \Sigma, \prec_f \rangle$ is a DAG. We will construct a truth predicate for Σ.

Let Σ_g be the set of all grounded points in $\langle \Sigma, \prec_f \rangle$. Note that Σ_g is conversely well-founded. Moreover, it is closed to $\prec_f$. So, by Theorem 2.5 and Lemma 3.4, there exists a subset Γ_g of Σ_g such that Γ_g is a truth predicate for Σ_g.

Let Σ^* be $\Sigma \setminus \Sigma_g$, i.e., the complement of Σ_g to Σ. We apply König's infinity lemma to find a truth predicate for Σ. Let $\Sigma^* = \{A_k \mid k \in \mathbb{N}\}$, and for any $k \in \mathbb{N}$, let $\Sigma^*_k = \{A_i \mid i < k\}$. For any $k \in \mathbb{N}$, we say a mapping s from $\{i \in \mathbb{N} \mid i < k\}$ to $\{0, 1\}$ is a *k-sequence*, if there is a set Γ_k such that (i) for any $A \in \Sigma^*_k \cup \Sigma_g$, $\Gamma_k(T\ulcorner A \urcorner) = \Gamma_k(A)$, (ii) $s(i) = \Gamma_k(A_i)$, and (iii) if $A \in \Sigma_g$, then $\Gamma_k(A) = \Gamma_g(A)$. k is the length of s. Note that Γ_k is a truth predicate for $\Sigma^*_k \cup \Sigma_g$. For convenience, it is called a "*truth witness*" to the k-sequence s. Let $\mathcal{T}$ be the set of k-sequences for all $k \in \mathbb{N}$, and let $<$ be a binary relation on $\mathcal{T}$ given by: $s < s'$, iff the length of s is less than that of s', and for all i less than the length of s, $s(i) = s'(i)$. See Figure 1 for an illustration of the idea.

[11] The notion of groundedness originates from Herzberger [12, p. 148]. The present one is the same as the one given by Halbach et al. [11, p. 193], who use the term "converse wellfounded" instead. See also [17, p. 694], [28, p. 122], [19, p. 168], [22, p. 748], and [2, p. 452] for other related but different notions of groundedness.

https://doi.org/10.1017/S1755020324000200 Published online by Cambridge University Press

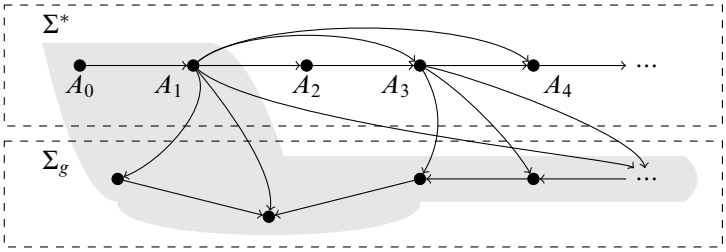


Figure 1. In this reference digraph, A_1 and A_3 are two weakly social sentences. By the shaded area, we indicate which sentences need to be considered in order to obtain a truth witness for a 2-sequence.

It can be easily verified that $\langle \mathcal{T}, < \rangle$ is a binary tree whose root is the empty sequence. For any $k \in \mathbb{N}$, since there are only finitely many social sentences in the set $\Sigma_k^* \cup \Sigma_g$, by Theorem 3.2, there is a truth predicate for it, which is a truth witness to some k-sequence. From this point, we can easily see that $\langle \mathcal{T}, < \rangle$ is an infinite tree. By König's infinity lemma, $\langle \mathcal{T}, < \rangle$ has a branch, namely τ. τ is a mapping from $\mathbb{N}$ to $\{0, 1\}$. Let Γ be the union of Γ_g and $\{A_k \in \Sigma^* \mid \tau(k) = 1\}$.

We prove that for any $A \in \Sigma$, $A \in \Gamma$, iff $\Gamma(A) = 1$. For convenience, for any $k \in \mathbb{N}$, we use τ_k for the restriction of τ to the set $\{i \in \mathbb{N} \mid i < k\}$. Fix $A \in \Sigma$, we consider the following four cases.

Case 1: $A \in \Gamma_g$. We want $\Gamma(A) = 1$. First, note that $\Gamma_g \subseteq \Sigma_g$ is a truth predicate for Σ_g, and so $\Gamma_g(A) = 1$. Besides, Σ_g is closed to $\prec_f$. Note Σ_g and Σ^* are disjoint. So, $\Gamma \cap f(A)$ equals $\Gamma_g \cap f(A)$. We get $\Gamma(A) = \Gamma_g(A) = 1$.

Case 2: $A = A_k \in \Sigma^*$ and $\tau(k) = 1$. Note that all social sentences of Σ are weak in the reference digraph $\langle \Sigma, \prec_f \rangle$, and Σ^* includes all ungrounded points in this digraph. We can find a number $n_k > k$ such that $f(A_k) \cap \Sigma^* \subseteq \Sigma_{n_k}$. By $\tau(k) = 1$, we know $\tau_{n_k}(k) = 1$. Since τ_{n_k} is an element of $\mathcal{T}$ with length n_k, we can find a corresponding truth witness Γ_{n_k} to τ_{n_k}. By $\tau_{n_k}(k) = 1$, it immediately follows $\Gamma_{n_k}(A_k) = 1$.

Our target is to prove $\Gamma(A_k) = \Gamma_{n_k}(A_k) = 1$. For this, it suffices to show $\Gamma_{n_k} \cap f(A_k) = \Gamma \cap f(A_k)$. Suppose $A \in \Gamma_{n_k} \cap f(A_k)$, then $\Gamma_{n_k}(A) = 1$. We consider two sub-cases. First, in case $A \in \Sigma_g$, then $\Gamma_g(A) = \Gamma_{n_k}(A) = 1$. But Γ_g is a truth predicate for Σ_g, so $A \in \Gamma_g$. Hence, $A \in \Gamma$. Second, in case $A \in \Sigma^*$, then by choice of n_k, we have known $f(A_k) \cap \Sigma^* \subseteq \Sigma_{n_k}$, and hence $A \in \Sigma_{n_k}$. That is, $A = A_i$ for some $i < n_k$. Then, from $A_i \in \Gamma_{n_k}$, it follows $\tau_{n_k}(A_i) = \Gamma_{n_k}(A_i) = 1$. Again, we get $A = A_i \in \Gamma$. To summarize, we obtain $\Gamma_{n_k} \cap f(A_k) \subseteq \Gamma \cap f(A_k)$.

It remains to prove $\Gamma \cap f(A_k) \subseteq \Gamma_{n_k} \cap f(A_k)$. Suppose $A \in \Gamma \cap f(A_k)$, then either (i) $A \in \Gamma_g$ or (ii) $A = A_i$ and $\tau(A_i) = 1$. In case (i), $\Gamma_{n_k}(A) = \Gamma_g(A) = \Gamma_g(T \ulcorner A \urcorner) = 1$, and so, $A \in \Gamma_{n_k}$. In case (ii), we notice $A = A_i \in f(A_k) \cap \Sigma^* \subseteq \Sigma_{n_k}$. Hence $i < n_k$. By $\tau(A_i) = 1$, we know $\Gamma_{n_k}(A_i) = \tau_{n_k}(A_i) = 1$. Again, we deduce $A_i \in \Gamma_{n_k}$. Consequently, we can conclude that $\Gamma \cap f(A_k)$ is a subset of $\Gamma_{n_k} \cap f(A_k)$.

Case 3: $A \in \Sigma_g \setminus \Gamma_g$. This is the dual of Case 1. We can show $\Gamma(A) = 0$. We leave the details to the reader.

Case 4: $A = A_k \in \Sigma^*$ and $\tau(k) = 0$. The proof is similar to the one in Case 2.

To sum up the above four cases, we can conclude that any sentence A belongs to Γ, iff $\Gamma(A) = 1$. Thus, Γ is a truth predicate for Σ. □

https://doi.org/10.1017/S1755020324000200 Published online by Cambridge University Press

Proof of Theorem 4.3. The proof is an almost verbatim version of the proof for Theorem 3.2 except that to the basis step, we apply Theorem 4.4 rather than Theorem 3.3. The details are omitted. □

The following corollary is a strengthening of Theorem 3.6. Their proofs are also similar.

THEOREM 4.5. *Suppose* $\langle \Sigma, \prec \rangle$ *is a reference digraph of* Σ. *If* Σ *is a non-self-referential and paradoxical set, then there is a ray* $A_0 \prec^* A_1 \prec^* A_2 \prec^* \ldots$ *in some sub-digraph* $\langle \Delta, \prec \rangle$ *of* $\langle \Sigma, \prec \rangle$ *such that for all natural number* n, A_n *is strongly social in* $\langle \Delta, \prec \rangle$.

Among all social sentences, by Theorem 4.3, only those strongly social sentences can determine the paradoxicality of a non-self-referential set of sentences. To take an example, for any number $n \geq 0$, let Y_n be a sentence such that

$$Y_n \leftrightarrow \neg T \ulcorner Y_{n+1} \urcorner \wedge \forall x \geq \bar{n}\, T \ulcorner \dot{x} = \dot{x} \urcorner$$

is provable in PA. Let Σ be the set of sentences Y_n plus $\bar{n} = \bar{n}$ $(n \geq 0)$, $\prec$ be a reference relation of $\mathcal{L}_T$. Then, every Y_n is weakly social in the reference digraph $\langle \Sigma, \prec \rangle$. Thus, Σ is not paradoxical even though it contains infinitely many social sentences.

Finally, it is worth pointing out that we can define the strongly social part of a digraph as we do in Definition 3.7. Then, we can prove that in any reference digraph of a non-self-referential and paradoxical set, the strongly social part must be conversely ill-founded. The proof is similar to that of Theorem 3.8, except that we appeal to Theorem 4.4 rather than Theorem 3.3.

§5. Dangerous Digraphs. In this section, we introduce the notion of the dangerous digraph and recapitulate the features of reference digraphs for paradoxes in purely graph-theoretic terms.

The notion of the dangerous digraph is originally given by Rabern et al. [22, p. 738] in a propositional language. It is extended into the first-order language of arithmetic with the truth predicate symbol T by Beringer & Schindler [2, p. 454].

DEFINITION 5.1. *A digraph is* dangerous, *if there is a paradoxical set of sentences such that one of its reference digraphs is isomorphic to this digraph*.

Note that if a set of sentences is locally finite, it must have a locally finite reference digraph. The following result is immediate from Corollary 2.7(1) and Theorem 3.3.

THEOREM 5.2.

(1) *Any finite and dangerous digraph contains at least a directed cycle.*
(2) *Any locally finite and dangerous digraph contains at least a directed cycle.*

Theorem 5.2 has been proved by Rabern et al. [22, pp. 751, 754]. See also [2, p. 474]. It is well-known that for any $n \geq 0$, the set $\{L_k \mid 0 \leq k \leq n\}$ with the provable equivalences $L_0 \leftrightarrow \neg T \ulcorner L_n \urcorner$ and $L_{k+1} \leftrightarrow T \ulcorner L_k \urcorner$ $(1 \leq k < n)$ is paradoxical. From this observation, we can also prove that any finite digraph containing a directed cycle must be dangerous. See also [22, pp. 750–751] for details.

Note that the reference digraph of the liar sentence is the minimal reflexive digraph $\langle \{0\}, = \rangle$. It is called the liar digraph by Rabern et al. [22, p. 738]. A corollary of the above theorem is that a finite digraph is dangerous, iff it contains a subdivision of the liar digraph as a sub-digraph (ibid., p. 751). The proof is straightforward. The reader can refer (ibid., p. 742) to the definition of subdivision.

https://doi.org/10.1017/S1755020324000200 Published online by Cambridge University Press

The first result of Theorem 5.2 is a particular case of the following. The latter one, in turn, is a graph-theoretic version of Theorem 2.5.[12]

THEOREM 5.3. *Every dangerous digraph is conversely ill-founded.*

Next, we turn to the dangerous digraphs containing no cycle. The following result is immediate from Corollary 2.7(2) and Theorem 3.2.

THEOREM 5.4.

(1) *Every dangerous DAG contains at least a ray.*
(2) *Every dangerous DAG contains infinitely many social points* (*i.e., points of infinite out-degree*).

To take an example, $\langle \mathbb{N}, < \rangle$ is (isomorphic to) a reference digraph for Yablo's paradox. It is one of the simplest dangerous DAGs. This linear structure is called the Yablo digraph by Rabern et al. [22, p. 738].

Theorem 5.4, as mentioned before, is first proved by Rabern et al. [22, pp. 750, 756] in an infinitary propositional language. The present version is proved based on Corollary 2.7(2) and Theorem 3.2, in which paradoxes are formulated in the first-order language of arithmetic with T. In this sense, our theorem can be seen as an extension of Rabern et al.'s [22, p. 756] corresponding result.

Theorem 3.6, as mentioned below its proof, is a strengthening of Corollary 2.7(2) and Theorem 3.2 in the sense that those infinitely many social sentences may simultaneously occur in a ray. We now reformulate Theorem 3.6 and Theorem 4.5 in terms of dangerous digraphs.

THEOREM 5.5. *Every dangerous DAG contains a sub-digraph in which there are infinitely many social points appearing in a ray. The same is true for strongly social points.*

We close this section with four variants of Yablo's paradox. In the following examples, the sentences are arranged in non-linear structures so that their reference digraphs are more and more complex than the Yablo digraph.

EXAMPLE 5.6 (Two-dimensional Yablo's paradox). *Define the set of sentences $Y_{m,n}$ for all $m, n \in \mathbb{N}$ such that $Y_{m,n}$ is the sentence saying that $Y_{i,j}$ is untrue for all $i, j \in \mathbb{N}$ with $m < i$ or $n < j$.*

EXAMPLE 5.7 (Nested two-dimensional Yablo's paradox). *Define the set of sentences $Y_{m,n}$ for all $m, n \in \mathbb{N}$ such that $Y_{m,n}$ is the sentence saying that $Y_{i,j}$ is untrue for all but finitely many $i, j \in \mathbb{N}$ with $m < i$ or $n < j$.*[13]

Recall that the Cantor tree (the infinite full binary tree) is $2^{<\omega}$, i.e., the set of the finite sequences of 0's and 1's, ordered by the relation $<$ such that a sequence is "less than" another iff the latter is a proper extension of the former. More formally, for any finite sequences s and t, $s < t$, iff $s \subsetneq t$.[14]

EXAMPLE 5.8 (Binary-tree Yablo's paradox). *Define the set of sentences Y_s for all $s \in 2^{<\omega}$ such that Y_s is the sentence saying that Y_t is untrue for all $t \in 2^{<\omega}$ with $s < t$.*

[12] An equivalent proposition of some different form has been obtained by Rabern et al. [22, Lemma 9, p. 749]. Rabern et al.'s version is proved in the setting of an infinitary propositional language.

[13] The corresponding one-dimensional one is initially given by Yablo [31, p. 144]. It is also called the $\exists\forall$-unwinding variant of Yablo's paradox. See also footnote 3.

[14] See [16, pp. 27ff] for more information about the tree.

https://doi.org/10.1017/S1755020324000200 Published online by Cambridge University Press

EXAMPLE 5.9 (Nested binary-tree Yablo's paradox). *Define the set of sentences Y_s for all $s \in 2^{<\omega}$ such that Y_s is the sentence saying that Y_t is untrue for infinitely many $t \in 2^{<\omega}$ with $s < t$, that is, for any $u \in 2^{<\omega}$ with $s < u$, there exists at least one $t \in 2^{<\omega}$ such that $u < t$ and Y_t is untrue.*

All of the above four examples are paradoxical. We verify the third one. Our verification is informal, but it is easy to transform it into a formal proof required by Definition 2.1. We denote the empty sequence by $\langle\rangle$. Assume $Y_{\langle\rangle}$ is true, then Y_s is false for all $s \in 2^{<\omega}$. In particular, $Y_{\langle 0\rangle}$ is false. At the same time, we also know Y_s is false for all s with $\langle 0\rangle < s$. Thus, $Y_{\langle 0\rangle}$ must be true, a contradiction. Hence, the assumption implies a contradiction, so $Y_{\langle\rangle}$ must be false. In that case, for some $s \in 2^{<\omega}$, Y_s is true. Similarly, we can derive a contradiction from the truth of Y_s. As a result, the set $\{Y_s \mid s \in 2^{<\omega}\}$ is paradoxical.

We can formulate the sentences in the above examples into the language $\mathcal{L}_T$. Again, we do it for the third one. First, note that the relation $<$ on $2^{<\omega}$ is a representable relation. Besides, there is a computable function δ from $\mathbb{N}$ to $2^{<\omega}$. The inverse mapping δ^{-1} encodes the finite sequences in $2^{<\omega}$. The number $\delta^{-1}(s)$ is the coding number of the sequence s.[15] Next, define a binary relation $\prec$ on $\mathbb{N}$: $n \prec m$, iff $\delta(n) < \delta(m)$. $\prec$ is a representable relation. Thus, by diagonalization, we can construct a sequence of sentences Y'_n such that $Y'_n \leftrightarrow \forall x\,(n \prec x \rightarrow \neg T \ulcorner Y'_{\dot{x}} \urcorner)$ is provable. Finally, for any $s \in 2^{<\omega}$, let $Y_s = Y'_{\delta^{-1}(s)}$. Then, the binary-tree Yablo's paradox is the set of sentences Y_s $(s \in 2^{<\omega})$.

Note that the Yablo digraph occurs in all reference digraphs of the above four examples as a finitary minor. Again, we only consider the set of sentences Y_s for all $s \in 2^{<\omega}$ in Example 4. We denote it by Σ. Suppose f is a dependence function such that $\prec_f$ is a reference relation on Σ, then $f(Y_s)$ is a co-finite subset of Σ. From this, we can find infinitely many strongly social sentences in some branch. This branch forms a sub-digraph of $\langle \Sigma, \prec_f \rangle$, which is isomorphic to the Yablo graph. We leave the details to the reader.

As mentioned above, a primary difference between the above four examples and Yablo's paradox is that their sentences are not arranged by a linear ordering. In the last two examples, the sentences are even arranged in a tree. Even if we try to spread sentences onto such a non-linear structure, so long as these sentences form a paradox, we can still find a ray in which there are infinitely many strongly social sentences.

§6. Concluding Remarks. This paper builds on the work by Rabern et al. [22] to explore the reference digraphs of the paradoxes. The starting point of our research is a result Rabern et al. establish in a specific infinitary propositional language: all dangerous acyclic digraphs contain infinitely many points with an infinite out-degree. We extend this result in the first-order arithmetic language with a primitive truth predicate. The version that we prove is that any reference digraph of a non-self-referential paradox contains infinitely many social sentences.

We strengthen the above result in two respects. On the one hand, among these social sentences, infinitely many appear in one ray. On the other hand, among these social sentences, infinitely many have infinitely many out-neighbors, none of which will eventually get to a sink.

[15] See, for instance, [7, p. 42].

https://doi.org/10.1017/S1755020324000200 Published online by Cambridge University Press

Based on the above observations, we finally discuss Beringer & Schindler's conjecture briefly. As mentioned in the introductory section, the non-trivial part of their conjecture states that all dangerous acyclic digraphs contain the Yablo digraph as a finitary minor. For a formal definition of the notion of a finitary, we refer the reader to Beringer & Schindler [2, p. 490]. For now, it is sufficient to understand that the essence of this conjecture lies in that starting from any dangerous acyclic digraph, we can obtain the Yablo digraph through a series of vertex deletions, edge deletions, or edge contractions (in any order).

By Theorem 5.5, we have established that a dangerous DAG must contain two of essential components of the Yablo graph: infinitely many strongly social points and a ray passing through these points. However, Theorem 5.5 alone is insufficient to address Beringer & Schindler's conjecture, as a complete Yablo graph cannot necessarily be constructed solely from these two components. To illustrate this, consider the digraph whose domain is the union of natural numbers and rational numbers of the form $n + \frac{1}{2^k}$ for all $n \in \mathbb{N}$ and $k \geq 1$. The binary relation of this digraph is the union of the successor relation on natural numbers and the relation $\left\{\langle n, n + \frac{1}{2^k}\rangle, \langle n + \frac{1}{2^k}, n + 1\rangle \mid n \in \mathbb{N}, k \geq 1\right\}$. In this digraph, the sequence of natural numbers forms a ray, with each natural number being a strongly social point. However, it is evident that the Yablo graph cannot be obtained from this digraph through vertex deletions, edge deletions, or edge contractions. In fact, it can be proved that such a digraph is not even dangerous.[16]

To prove Beringer & Schindler's conjecture, we need to further strengthen Theorem 5.5. For this, we introduce the notion of domination between a point and a ray. We say a point dominates a ray if there are infinitely many paths starting from that point and terminating at the ray, such that these paths are pairwise disjoint except for the common starting point. Moreover, no point between the starting point and the terminating point (exclusive) in each of these paths occurs in the ray. A sufficient condition for Beringer & Schindler's conjecture is that every dangerous DAG contains a sub-digraph in which there are infinitely many (strongly) social points dominating a ray.[17] Seeing this point, we would like to say that the information we extract from Theorem 3.2 and Lemma 3.5 is not sufficiently comprehensive, despite the significant

[16] This example is originally given by Walicki [26, p. 474] with a slight variation.

[17] The notion of a point dominating a ray comes from Walicki [26, p. 476]. The author is grateful to an anonymous reviewer of a previous version of this manuscript for bringing this work to his attention and highlighting the application of kernel theory to paradoxes. Roughly speaking, a kernel S is a set of points in a digraph such that any point in S has no (out-)neighbor in S and any point outside S has at least an out-neighbor in S. The notion of the kernel can be traced back to the notion of solution for binary relations introduced by Von Neumann & Morgenstern [24]. As far as we know, it is Cook [5, p. 773] who first applies kernel theory to analyze a specific class of paradoxes obtained by "unwinding". Recently, Cook's ideas have been further developed in works such as Cook [6], Dyrkolbotn & Walicki [8], Walicki [25], Walicki & Dyrkolbotn [27], and Walicki [26]. In particular, the works of Walicki et al. have made significant contributions to the application of kernel theory to all semantic paradoxes by transforming formulas into graph normal forms in an infinitary propositional language. Additionally, it is worth noting that in Walicki [26, p. 476], a digraph is defined as "safe" if it contains neither odd cycles nor any ray that is dominated by infinitely many points within that ray. These unsafe digraphs are closely related to the reference digraph of paradoxes in graph normal forms. Further investigation is needed to explore the relationship between dangerous digraphs and Walicki's unsafe digraphs.

https://doi.org/10.1017/S1755020324000200 Published online by Cambridge University Press

result of Theorem 3.6 obtained from these two statements. There remains a gap between Theorem 3.6 and the aforementioned condition. Further investigation is required to bridge this gap.

Acknowledgments. I would like to thank an anonymous referee of this journal for her/his helpful comments and suggestions. Additionally, I am grateful to an anonymous referee of an earlier version of this paper for pointing out a significant gap in an attempt to prove Beringer & Schindler's conjecture.

Funding. This research is supported by National Social Science Fund of China (A Truth-theoretical Study on Limits of Artificial Intelligence, No. 24AZX018).

BIBLIOGRAPHY

[1] Beringer, T., & Schindler, T. (2016). Reference graphs and semantic paradox. In ArazimP., & Dančák, M. , editors. *Logica Yearbook 2015*, London: College Publications, pp. 1–15.

[2] ———. (2017). A graph-theoretic analysis of the semantic paradoxes. *The Bulletin of Symbolic Logic*, **23**(4), 442–492.

[3] Boolos, G. (1993). *Logic of Provability*. Cambridge: Cambridge University Press.

[4] Butler, J. M. (2018). An entirely non-self-referential Yabloesque paradox. *Synthese*, **195**(11), 5007–5019.

[5] Cook, R. T. (2004). Patterns of paradox. *Journal of Symbolic Logic*, **69**(3), 767–774.

[6] ———. (2014). *The Yablo Paradox: An Essay on Circularity*. Oxford: Oxford University Press.

[7] Cutland, N. J. (1980). *Computability: An Introduction to Recursive Function Theory*. Cambridge: Cambridge University Press.

[8] Dyrkolbotn, S., & Walicki, M. (2014). Propositional discourse logic. *Synthese*, **191**(5), 863–899.

[9] Feferman, S. (1960). Arithmetization of metamathematics in a general setting. *Fundamenta Mathematicae*, **49**(1), 35–92.

[10] Gupta, A. (1982). Truth and paradox. *Journal of Philosophical Logic*, **11**(1), 1–60.

[11] Halbach, V., Leitgeb, H., & Welch, P. (2003). Possible-worlds semantics for modal notions conceived as predicates. *Journal of Philosophical Logic*, **32**(2), 179–223.

[12] Herzberger, H. G. (1970). Paradoxes of grounding in semantics. *The Journal of Philosophy*, **67**(6), 145–167.

[13] ———. (1982). Notes on naive semantics. *Journal of Philosophical Logic*, **11**(1), 61–102.

[14] Hsiung, M. (2020). What paradoxes depend on. *Synthese*, **197**(2), 887–913.

[15] ———. (2021). Unwinding modal paradoxes on digraphs. *Journal of Philosophical Logic*, **50**(2), 319–362.

[16] Just, W., & Weese, M. (1997). *Discovering modern set theory II: set-theoretic tools for every mathematician. Graduate Studies in Mathematics Volume 18*. Providence, RI: American Mathematical Society.

https://doi.org/10.1017/S1755020324000200 Published online by Cambridge University Press

[17] Kripke, S. A. (1975). Outline of a theory of truth. *Journal of Philosophy*, **72**(19), 690–712.

[18] Leitgeb, H. (2002). What is a self-referential sentence? Critical remarks on the alleged (non-)circularity of Yablo's paradox. *Logique et Analyse*, **45**(177-178), 3–14.

[19] ———. (2005). What truth depends on. *Journal of Philosophical Logic*, **34**(2), 155–192.

[20] McGee, V. (1985). How truthlike can a predicate be? A negative result. *Journal of Philosophical Logic*, **14**(4), 399–410.

[21] Priest, G. (1997). Yablo's paradox. *Analysis*, **57**(4), 236–242.

[22] Rabern, L., Rabern, B., & Macauley, M. (2013). Dangerous reference graphs and semantic paradoxes. *Journal of Philosophical Logic*, **42**(5), 727–765.

[23] Schlenker, P. (2007). The elimination of self-reference: generalized Yablo-series and the theory of truth. *Journal of Philosophical Logic*, **36**(3), 251–307.

[24] Von Neumann, J., & Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.

[25] Walicki, M. (2017). Resolving infinitary paradoxes. *Journal of Symbolic Logic*, **82**(2), 709–723.

[26] ———. (2019). Kernels of digraphs with finitely many ends. *Discrete Mathematics*, **342**(2), 473–486.

[27] Walicki, M., & Dyrkolbotn, S. (2017). Diagnosing paradoxes. Proceeding of TARK *2017*.

[28] Yablo, S. (1982). Grounding, dependence, and paradox. *Journal of Philosophical Logic*, **11**(1), 117–137.

[29] ———. (1985). Truth and reflection. *Journal of Philosophical Logic*, **14**(3), 297–349.

[30] ———. (1993). Paradox without self-reference. *Analysis*, **53**(4), 251–252.

[31] ———. (2006). Circularity and paradox. In T. Bolander, V. F. Hendricks, & S. A. Pedersen, editors. *Self-Reference*, Stanford: CSLI Publications, pp. 139–157.

SCHOOL OF PHILOSOPHY AND SOCIAL DEVELOPMENT
SOUTH CHINA NORMAL UNIVERSITY
GUANGZHOU 510631
PEOPLE'S REPUBLIC OF CHINA
E-mail: mingshone@163.com

https://doi.org/10.1017/S1755020324000200 Published online by Cambridge University Press