Notre Dame Journal of Formal Logic
Volume 35, Number 1, Winter 1994

Reasoning with Sentences and Diagrams

ERIC HAMMER

Abstract A formal system is studied having both sentences and diagrams as well-formed representations. Proofs in the system allow inference back and forth between sentences and diagrams, as well as between diagrams and diagrams, and between sentences and sentences. This sort of heterogeneous system is of interest because external representations other than linguistic ones occur commonly in actual reasoning in conjunction with language. Syntax, semantics, and rules of inference for the system are given and it is shown to be sound and complete.

1 Introduction The representations used in actual inference come in many different forms. Besides language, there are diagrams, charts, graphs, tables, etc. Moreover,

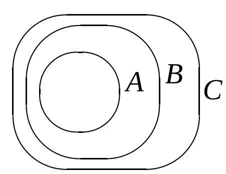


Figure 1: An Euler diagram.

often two or more different types of representations are used in the very same proof or reasoning task with great success. To give a simple example, in actual practice one would routinely assert that the two sentences "All A's are B's" and "All B's are C's" together express the same claim as Euler diagram in Figure 1. Similarly, it would be typical to state that it can be immediately seen from the diagram that all A's are C's, that this can be "read off" of the diagram or "inferred" from it. Beyond such simple examples, diagrams occur in very complicated mathematical proofs for such purposes as revealing the overall structure of a proof, clarifying the main construction of a proof, etc. They are used in many other types of complicated reasoning tasks as well. So in informal practice at any rate, inference between sentences and diagrams and assertions of synonymy between diagrams and sentences are a matter of course.

Received July 7, 1992; revised August 28, 1993

Perhaps it is common wisdom to say that while diagrams do appear at the external level of the proofs people actually give, still they are not an essential part of these proofs. Rather, it is said that they are a heuristic tool or psychological aid, that the real proof is best modeled as some finite sequence of sentences. Nevertheless, it is not at all clear that this is a view that should be defended at any cost. Such things as the structure of the proofs given in informal practice, the length of the proofs given, the types of rules applied and in what order, the simplicity of the rules used, etc. have all been considered to be important factors in the construction of logical systems. For example, natural deduction systems are motivated by a desire to accurately model the way proofs are actually structured, such as the use of temporary assumptions in a proof, the method of breaking into cases, the use of proof by contradiction, etc. Likewise, a second-order logic rather than a first-order logic might be taken to model some inference practice simply because it appears to more adequately reflect the type of inferences actually made, as illustrated by Boolos in [1]. Mere extensional accuracy is not the only demand made of a mathematical model of a given inference practice. It is only one among many.

This being the case, there seems no reason why one shouldn't consider the type of representation used in actual proofs to be a legitimate motivation in the construction of logics. Since diagrams are used in key places in mathematical proofs, they should appear in the same key places in formal accounts of those types of proofs. Likewise, since inferences are commonly made from diagrams to sentences and vice versa, such inferences should also be duly analyzed from a logical point of view. The present paper attempts to do this for a simple case.

A "heterogeneous" inference system will be analyzed, heterogeneous in having as representations both sentences and diagrams and in allowing inference between the two types of representations. The system is based on Shin's work in [3] and [4]. Shin studies the syntax, semantics, and model theory of two purely diagrammatic systems of Venn diagrams. In addition to the Venn diagrams studied by Shin, the present system allows information to be represented by means of first-order sentences. Rules of inference allow one to make inferences from sentences to sentences, from sentences to diagrams, from diagrams to diagrams, and from diagrams to sentences. While the system is a fairly simple example of a heterogenous logic, it will hopefully serve to illustrate that such systems are a legitimate topic for logical analysis and also to raise some topics of concern for them.

The syntax of the system's well formed representations, both diagrams and sentences, is described first. Semantics are given which encompass both types of representations, thereby allowing for meaningful interaction between the two. Rules of transformation are given, some of which are standard first-order rules, some of which are purely diagrammatic, and some of which are heterogeneous. The rules are shown to be sound and complete with respect to the given semantics. Besides the inference rules, the diagrams themselves of the system are heterogeneous, having both diagrammatic and linguistic elements. The question therefore arises as to whether a diagram is a consequence of another in virtue of its "diagrammatic features" or whether merely in virtue of its "linguistic features." The question also arises as to whether the purely diagrammatic rules of the system are complete with respect to this notion of "diagrammatic consequence." These matters are discussed in the final section.

There are many potentially interesting topics for further study concerning the formal properties of heterogeneous logics. For example, the presence of two or more different types of representations in such a logic suggests various comparisons of the different components. Does one subsume the other with respect to expressive power? Is one more computationally efficient than the other? Are there any advantages (such as length of proof) with proofs involving both (or all three, etc.) types of representations as opposed to proofs using just one type? Is one type of representation more closely tied to the semantics than the other? Are the different types of representations better suited to different types of expressive tasks? What sorts of claims is each component best designed to make? Is each component of the system complete?

2 Well-Formed Representations The primitive diagrammatic objects of the system include both linguistic and diagrammatic objects. Among the linguistic primitives are the basic symbols of first-order logic plus a lambda operator used to bind free variables in formulas:

1. Logical constants: $\forall$, $\exists$, $\rightarrow$, $\vee$, $\wedge$, $\neg$, and λ
2. Constant symbols: $a, b, c, a_1, b_1, c_1, \ldots$
3. Variable symbols: $x, y, z, x_1, y_1, z_1, \ldots$
4. Predicate symbols of each arity: $P, Q, R, P_1, Q_1, R_1, \ldots$

The diagrammatic primitives of the system consist of the "Rectangle," "Closed Curve," "Shading," "Line" and "X" as shown below:

The terms and well-formed formulas (wffs) of the system are formed in the usual way. In addition, the system has set terms formed by abstracting over the free variable in a wff. These set terms are used to tag the closed curves of Venn diagrams. In other words, the set of "set terms" is the smallest class satisfying the following condition: $\lambda x \varphi$ is a set term whenever x is a variable and φ is a wff having at most the variable x occurring free. For example, $\lambda x\, Pilot(x)$ is a set term and will be interpreted in a model as the extension of the predicate $Pilot$, i.e., as the set of pilots.

An "X-sequence" is a finite number of X's connected by lines into a chain. For example, "$\otimes - \otimes - \otimes - \otimes$" is an X-sequence. Likewise, for any constant symbol b, a "b-sequence" is a finite number of tokens of b connected by lines into a chain, such as "$b - b - b$."

Definition 2.1 The set of "well-formed diagrams" (wfds) is the smallest class satisfying the following four conditions:

1. Any rectangle is a wfd.
2. If D is a wfd and C is a closed curve labelled by exactly one set term not occurring in D, then the diagram obtained by adding C to D in accordance with the partial overlapping rule is a wfd. The partial overlapping rule requires that C intersect each enclosed region of D exactly once, and that it overlaps only part of each enclosed region.

3. If D is a wfd and b is any constant symbol, then the diagram obtained by adding a b-sequence or X-sequence to D is a wfd, provided that every link of the sequence falls entirely within the rectangle and does not contact any border of a closed curve of D.
4. If D is a wfd, then the diagram obtained by shading some enclosed area of D is a wfd, provided that the shading is entirely bounded by parts of closed curves and the rectangle.

A "well-formed representation" (wfr) is any wff or wfd. A "closed wfr" is any sentence or wfd. The two diagrams in Figure 2, for example, are well-formed.

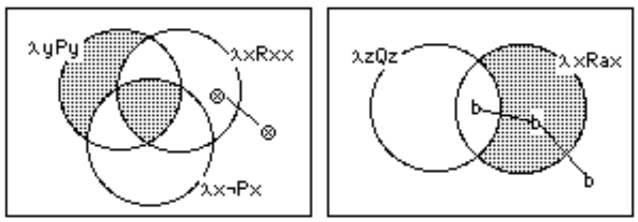


Figure 2: Two well-formed diagrams.

Intuitively, the region enclosed by a closed curve represents the set indicated by its label. The region of overlap of two regions represents the intersection of the two sets represented by the two regions, and so on. The shading of a region is an assertion that the set represented by that region is empty. Similarly, an X-sequence asserts non-emptiness of the set represented, and a b-sequence asserts that b is a member of the set represented. For example, in the right-hand diagram in Figure 2, the left-hand closed curve represents the Q's, and the right-hand curve represents those objects x such that a bears R to x. The diagram asserts (by the shading) that there is no object x which is non-Q such that a bears R to x, and (by the b-sequence) that b is such that either a bears R to b or else both b is not Q and a does not bear R to b.

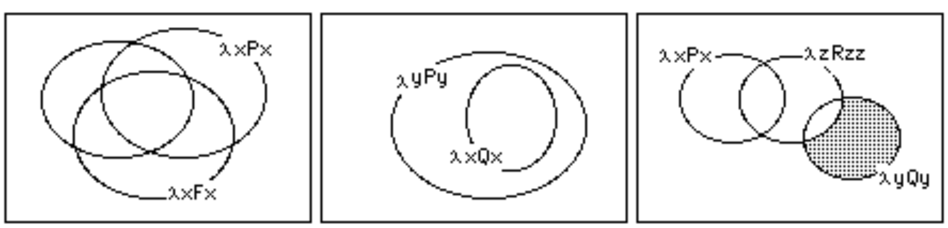


Figure 3: Some non well-formed diagrams.

On the other hand, the three diagrams in Figure 3 are not well-formed. The first one has a closed curve not labelled by a set term. The second and third ones violate the partial overlapping rule.

The relevant syntactic units of a wfd are its "regions" which are defined by its rectangle and closed curves. Regions are defined as follows: A "basic region" of a wfd D is any region enclosed by a closed curve occurring in D or enclosed by the rectangle of D. Thus, in a wfd having n closed curves there are $n + 1$ basic regions. The "regions" of D are then defined inductively as follows:

1. If r and s are regions of D, then the combined region consisting of r together with s is a region (denoted by "$r \cup s$");
2. If r and s are regions of D that overlap, then the region composed of the area of overlap of r and s is a region (denoted by "$r \cap s$");
3. If r and s are regions of D and s is a proper part of r, then the region that is part of r but not part of s is a region (denoted by "$r - s$");
4. If r is a region of D, then the region within the rectangle of D but outside of r is a region (denoted by "$\overline{r}$").

Finally, a "minimal region" is any region which has no other region as a proper part, and a "subregion" of a region r is any region s which is a part of r. If r is a subregion of s, s is also said to "contain" r. A wfd having n closed curves will have 2^n minimal regions and $2^{2^n} - 1$ regions.

Note that the operations $\cup$, $\cap$, etc. in the definition of regions are operations on parts of diagrams, i.e., on syntactic objects, not on sets. The same operation symbols will be also be used with their usual set-theoretic meaning, but it will be clear from context which operation is intended.

3 Correspondence Between Regions and Set Terms While the logical connections between different sentences are often subjected to analysis, the corresponding connections between sentence and diagram are not. However, with its requirement that every closed curve be labelled by exactly one set term, the present system has all the necessary apparatus for making the logical connections between sentences and diagrams explicit.

The set terms that tag the closed curves of wfds will tell us what sets those closed curves are meant to represent. For example, if $\lambda x Pilot(x)$ tags a curve enclosing region r, then r represents the set of objects that are pilots. If $\lambda x(P(x) \vee Q(x))$ tags the curve, then r represents the objects that are either P or Q. And so on. Beyond these basic regions, we are also interested in the set represented by the overlap of two closed curves and the region enclosed by either curve, as well as the sets represented by more complex regions. For example, if $\lambda x P(x)$ and $\lambda x R(x, x)$ tag two closed curves, we want to be able to conclude that the region of overlap represents those objects that have the property P and are also R-related to themselves. Similarly, we want to be able to conclude that the region enclosed by the first but not the second closed curve represents those objects that are P but are not R-related to themselves, that the region enclosed by the rectangle but outside of the first closed curve represents those objects that are not P, etc.

To precisely capture this intuition, it is necessary to syntactically connect the basic syntactic elements of wfds (regions) to set terms and thereby to language. With this in mind, we will define a syntactic "correspondence relation," in symbols "$\cong$," which will hold between the regions of diagrams and various set terms derived from those occurring in them. This correspondence relation will hold, for example, between the region enclosed by a curve and its tag, and also between more complex regions and set terms composed appropriately from tags of D. The correspondence relation will determine the set a given region gets interpreted as in a particular model. It will play the role of intermediary between diagrams and sentences.

Definition 3.1 Let D be an arbitrary diagram and x, y, and z be arbitrary variables. Then the relation $\cong$ is defined as the smallest relation satisfying the two conditions:

1. If $\lambda x\varphi(x)$ tags a closed curve of D and r is the region enclosed by that closed curve, then $r \cong \lambda y\varphi(y)$ provided y is free in $\varphi(y)$.
2. If $r \cong \lambda x\varphi(x)$, $s \cong \lambda y\psi(y)$, and z is free in both $\varphi(z)$ and $\psi(z)$, then: $\overline{r} \cong \lambda z\neg\varphi(z)$, $r \cap s \cong \lambda z(\varphi(z) \wedge \psi(z))$, $r \cup s \cong \lambda z(\varphi(z) \vee \psi(z))$, $r - s \cong \lambda z(\varphi(z) \wedge \neg\psi(z))$, and $\overline{r} \cup s \cong \lambda z(\varphi(z) \to \psi(z))$.

To illustrate the correspondence relation, consider the diagram D in Figure 4. Let r be the region within the rectangle but outside all of its closed curves. Then $r \cong \lambda x\neg(P(x) \vee Q(x) \vee \forall y P(y))$, $r \cong \lambda y(\neg P(y) \wedge \neg Q(y) \wedge \neg\forall y P(y))$, and so on. Let s be the region of overlap of the upper two curves. Then $s \cong \lambda z(P(z) \wedge Q(z))$, $s \cong \lambda x\neg(P(x) \to \neg Q(x))$, etc. Let r' be the region enclosed by the upper right-hand closed curve. Then $r' \cong \lambda x Q(x)$, $r' \cong \lambda x\neg\neg Q(x)$, and so on. It is easy to verify

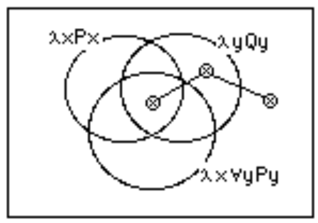


Figure 4: A diagram.

that there is an effective procedure for determining whether any given region and set term stand in the correspondence relation.

4 Counterpart Relation Between Regions It is convenient to define a "counterpart" relation that holds between regions of different diagrams signifying that the two regions are meant to represent the same set. For example, if two closed curves are both tagged with the same set term, then the two regions enclosed by the curves are intended to represent the same set. Further, the two regions falling outside the two closed curves also are intended to represent the same set. Similarly, if regions r of diagram D and r' of D' represent the same set, and s of D and s' of D' represent the same set, then the region of overlap of r and s represents the same set as the region of overlap of r' and s', and likewise for the operations $\cup$ and - on regions. More precisely:

Definition 4.1 The "counterpart relation" is defined inductively as the smallest binary relation on regions of diagrams such that for any two diagrams D having regions r and s and D' having regions r' and s':

1. If r and r' are regions enclosed by closed curves tagged with the same set term, then r and r' are counterparts, and
2. If r and r' are counterparts and s and s' are counterparts, then so are: $\overline{r}$ and $\overline{s}$; $r \cup s$ and $r' \cup s'$; $r \cap s$ and $r' \cap s'$; and $r - s$ and $r' - s'$.

5 Models and Truth in a Model Models for the system will be extentions of standard first-order models. As with first-order models, they will assign objects and sets of tuples to terms and predicate symbols. They will also assign sets to regions of diagrams in a way that respects the terms that tag closed curves. The sets assigned to more complicated regions will depend systematically on the sets assigned to simpler regions.

"First-order model," "assignment function of values to variables," "satisfaction of a wff φ by an assignment function v in a model M" (in symbols $M \models_{\textit{first-order}} \varphi[v]$), and "truth of a sentence in a model" (in symbols $M \models_{\textit{first-order}} \varphi$) are all defined in the usual way. The class of first-order models is now used to define "extended first-order models," which are first-order models extended in a natural way to assign subsets of the domain to each set term:

Definition 5.1 A pair (U, I') is an "extended first-order model" provided that there is some first order model (U, I) such that: (1) The interpretation function I' extends I, and (2) For every set term $\lambda x\varphi$, $I'(\lambda x\varphi) = \{d \in U : (U, I) \models \varphi[v]$ for some assignment function v on U assigning d to the variable $x\}$.

Intuitively, this says simply that I' must assign to a set term $\lambda x\varphi$ the set of all objects in the domain satisfying the formula φ in (U, I).

Proposition 5.2 *For every extended first-order model M', there is a unique first-order model M such that M' extends M. Furthermore, every first-order model can be extended to an extended first-order model.*

The extended first-order models are now extended to full-fledged models. These models extend the extended first-order models by assigning sets to regions of diagrams in accordance with the set terms labelling the closed curves of the diagrams. Thus:

Definition 5.3 A pair (U, I') is a "model" provided that there is an extended first-order model (U, I) and function F from regions of diagrams into the powerset of U such that $I' = I \cup F$, and for every wfd D:

1. If r is a region of D enclosed by a closed curve tagged with $\lambda x\varphi$, then $F(r) = I(\lambda x\varphi)$,
2. If r is the region of D enclosed by the rectangle, then $I(r) = U$, and
3. If r and s are regions of D, then $F(\bar{r}) = U - I(r)$, $F(r \cup s) = I(r) \cup I(s)$, $F(r \cap s) = I(r) \cap I(s)$, and $F(r - s) = I(r) - I(s)$.

Proposition 5.4 *For every model M'' there is exactly one extended first order model M' and exactly one first-order model M such that M'' extends M' and M' extends M. Furthermore, every extended first-order model can be extended to a model.*

Proof. Let (U, I) be an extended first-order model. Then define an extention I' of I to be $I \cup F$ where F is defined inductively such that for any diagram D,

1. If r is the region enclosed by a closed curve of D tagged with $\lambda x\varphi$, then $F(r) = I(\lambda x\varphi)$.
2. If r is the region enclosed by the rectangle, then $F(r) = U$.
3. If r is a minimal region and $s_1, \ldots, s_n$ are all the regions of closed curves that include r plus the region of the rectangle, and $t_1, \ldots, t_m$ are all the regions enclosed by closed curves that r falls outside of, then $F(r) = (F(s_1) \cap \ldots \cap F(s_n)) - (F(t_1) \cup \ldots \cup F(t_m))$.
4. If r is a region and $s_1, \ldots, s_n$ are the minimal regions constituting r, then $F(r) = F(s_1) \cup \ldots \cup F(s_n)$.

It is now routine to verify that F has the desired properties, and thus that (U, I') is a model that extends (U, I), finishing the proof of the proposition.

We are now prepared to define the notions of "truth of a wfd D in a model M" (in symbols $M \models D$) and "truth of a sentence φ in a model M" (in symbols $M \models \varphi$).

Definition 5.5 Let (U, I) be a model, D be a wfd, and φ be a sentence. Then:

1. $(U, I) \models D$ if and only if for every region r of D: if r is shaded then $I(r)$ is empty; if r has an X-sequence then $I(r)$ is non-empty; and if r has a b-sequence, then $I(b) \in I(r)$, and
2. $(U, I) \models \varphi$ if and only if $(U, I') \models_{\textit{first-order}} \varphi$, where (U, I') is the first-order fragment of (U, I).

A closed wfr δ is a "logical consequence" of a set of closed wfrs Δ (in symbols $\Delta \models \delta$) provided that every model M such that every member of Δ is true in M is such that δ is true in M. A sentence φ is a "first-order consequence" of a set of sentences Γ (in symbols $\Gamma \models_{\textit{first-order}} \varphi$) provided that every first-order model M such that every member of Γ is true in M is such that φ is true in M.

6 Interpretation Lemma The interpretation lemma plays a key role in both the soundness and completeness theorems. It systematically links up the interpretation of regions with the interpretation of set terms via the correspondence relation. Intuitively, it states that a region is interpreted in a model as the same set as any set term corresponding to it.

Theorem 6.1 Interpretation Lemma *Let (U, I) be a model, $\lambda x\varphi$ be a set term, and r be a region of D. Then, if $r \cong \lambda x\varphi$, then $I(r) = I(\lambda x\varphi)$.*

Proof. The proof is by induction on φ . For the base case, let φ be atomic. Then by the construction of $\cong$, r must be the region enclosed by some closed curve of D which $\lambda x\varphi$ tags. So by the definition of model we have that $I(r) = I(\lambda x\varphi)$. For the induction steps, suppose the lemma holds for every set term of lesser complexity than $\lambda x\varphi$. There are several cases:

1. $r \cong \lambda x(\psi \wedge \pi)$ where $\psi \wedge \pi$ is φ . We can assume that r is not a region enclosed by a closed curve which $\lambda x\varphi$ tags, since in that case the result follows immediately from the definition of model. By the definition of $\cong$ there must be regions s and t of D such that r is $s \cap t$, $s \cong \lambda x\psi$, and $t \cong \lambda x\pi$. Therefore, $I(r) = I(s \cap t)$ since r is $s \cap t$, which equals $I(s) \cap I(t)$ by the definition of model, which is $I(\lambda x\psi) \cap I(\lambda x\pi)$ by the induction hypothesis, which is $\{d \in U : (U, I) \models \psi[v]$ for some assignment function v assigning d to $x\} \cap \{d \in U : (U, I) \models \pi[v]$ for some assignment function v assigning d to $x\}$ by the definition of model, which equals $\{d \in U : (U, I) \models (\psi \wedge \pi)[v]$ for some assignment function v assigning d to $x\}$ by the definition of satisfaction, which is $\{d \in U : (U, I) \models \varphi[v]$ for some assignment function v assigning d to $x\}$ since $\psi \wedge \pi$ is φ, which is $I(\lambda x\varphi)$ by the definition of model.
2. $r \cong \lambda x\forall y\psi(x)$ where $\forall y\psi$ is φ . By the construction of $\cong$, r must be the region enclosed by some closed curve of D that is tagged by $\lambda z\forall y\psi(z)$ for some z that is free in $\psi(z)$. Thus, by definition of model, $I(r) = I(\lambda x\varphi)$. The cases for the other connectives are proved in a similar fashion, concluding the proof of the interpretation lemma.

Lemma 6.2 Counterpart Lemma *Let (U, I) be a model and let r and s be regions of diagrams D and D'. Then if r and s are counterparts, then $I(r) = I(s)$.*

Proof. Since every closed curve of every diagram must be tagged with a set term, and since r and r' are counterparts, there is a wff φ such that $r \cong \lambda x\varphi$ and $r' \cong \lambda x\varphi$. By the interpretation lemma, then, $I(r) = I(\lambda x\varphi) = I(r')$, as desired.

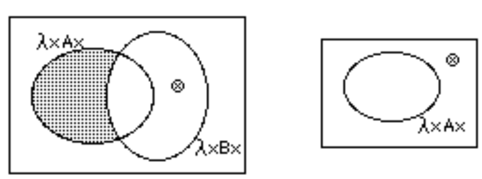
7 Rules of Inference The rules of the system are of three sorts. Diagrammatic rules allow one to infer a diagram from other diagrams. Heterogeneous rules allow one to infer a first-order sentence from a diagram, or to infer a new diagram from a first- order sentence and a diagram. First-order rules allow one to infer a sentence from other sentences.

A closed wfo φ is "provable" from a set Γ of closed wfo's (in symbols $\Gamma \vdash \varphi$) if and only if there is a finite sequence of sentences and diagrams, each of which is either a member of Γ, an axiom, or obtainable from earlier members of the sequence by one of the rules of inference. A sentence φ is "first-order provable" (in symbols $\Gamma \vdash_{\textit{first-order}} \varphi$) from a set Γ of sentences if and only if there is a finite sequence of sentences, each of which is either a member of Γ, a first-order axiom, or obtainable from earlier members of the sequence by one of the first-order rules of inference.

7.1 Diagrammatic Rules

Setup: A wfd with no shading or sequences may be asserted at any line of a proof.

Erasure: D' is obtainable from D by this rule if and only if D' results from either erasing a closed curve of D, erasing the shading of some region of D, or erasing an entire X-sequence of D. If it is a curve that is erased, any shading that would fill only part of some minimal region upon the erasure of the curve must also be erased from that minimal region. For example, the right-hand diagram below follows from the left-hand one by the erasure of a closed curve:



Extention of a Sequence: D' is obtainable from D by this rule if and only if D' results from D by the addition of extra links to some sequence of D.

Erasure of Links: D' is obtainable from D by this rule if and only if D' results from D by the erasure of links of a sequence falling in shaded regions, provided the remaining links are reconnected.

Unification of Diagrams: D'' is obtainable from D and D' by this rule if and only if:

Every region of D'' is the counterpart of a region of either D or D'. Conversely, every region of either D or D' is the counterpart of a region of D''.

If any region of D'' is shaded (has an X-sequence), it has a counterpart in either D or D' which also is shaded (has an X-sequence). Conversely, if a region of either D or D' is shaded (has an X-sequence), it has a counterpart in D'' that is shaded (has an X- sequence).

Non-Emptiness: D' is obtainable from D by this rule if and only if D' occurs from D by the addition of a sequence such that some link of the sequence falls into every minimal region of D.

7.2 Heterogeneous Rules The heterogeneous rules are of two sorts. Observe rules allow one to infer a sentence from a diagram, while apply rules allow one to apply the information expressed by a sentence to a diagram.

$\forall$-Apply: D' is obtainable from the sentence $\forall x\varphi$ and the diagram D by this rule if and only if there is some region r of D such that $r \cong \lambda x\varphi$ and D' results from D by the shading of any subregion of $\overline{r}$ or the addition of any sequence to r.

$\forall$-Observe: The sentence $\forall x\varphi$ is obtainable from D by this rule if and only if there is some region r of D such that $r \cong \lambda x\varphi$ and $\overline{r}$ is shaded.

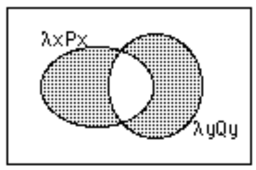


Figure 5: A diagram.

For example, $\forall x((P(x) \wedge Q(x)) \vee (\neg P(x) \wedge \neg Q(x)))$ is obtainable from the diagram in Figure 5, since the non-shaded region corresponds to $\lambda x((P(x) \wedge Q(x)) \vee (\neg P(x) \wedge \neg Q(x)))$. The sentence $\forall x((\neg P(x) \wedge \neg Q(x)) \vee (P(x) \wedge Q(x)))$ is also obtainable from the diagram by the same rule, as is $\forall x((P(x) \rightarrow Q(x)) \wedge (\neg P(x) \rightarrow \neg Q(x)))$ and many others. All of these are inferences one would be inclined to make in informal practice. So the motivation behind using the correspondence relation to state these rules is to have relatively powerful but natural heterogeneous rules which correspond relatively closely to informal practice.

$\exists$-Apply: D' is obtainable from the sentence $\exists x\varphi$ and the diagram D by this rule if and only if there is a region r of D such that $r \cong \lambda x\varphi$ and D' results from D by the addition of an X-sequence to some region containing r.

$\exists$-Observe: The sentence $\exists x\varphi$ is obtainable from D by this rule if and only if there is a region r of D such that $r \cong \lambda x\varphi$ and either some subregion of r contains a sequence or else $\overline{r}$ is shaded.

Constant-Apply: D' is obtainable from the sentence $\varphi(a)$ and the diagram D by this rule if and only if there is a region r of D such that $r \cong \lambda x\varphi(x)$ and D' results from D by the addition of an a-sequence to some region containing r.

Constant-Observe: The sentence $\varphi(a)$ is obtainable from D by this rule if and only if there is a region r of D such that $r \cong \lambda x\varphi(x)$ and either some subregion of r contains an a-sequence or else $\overline{r}$ is shaded.

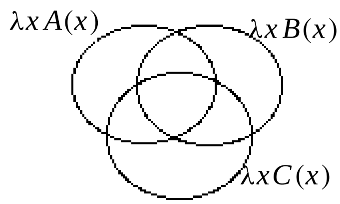
Inconsistent Information: Any closed wfo π is obtainable from D by this rule if and only if there is a shaded region r of D with a sequence in one of its subregions.

(1) $\forall x(Bx \rightarrow ((Ax \wedge \neg Cx) \vee (Cx \wedge \neg Ax)))$ Premise

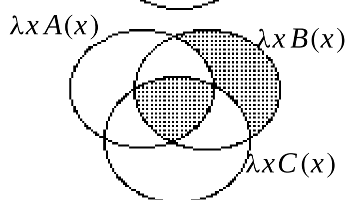
(2) $\forall x((Ax \vee Cx) \rightarrow (\neg Bx \vee (Bx \wedge Cx)))$ Premise

(3) $\forall y((Cy \vee By) \wedge \neg(Cy \wedge By \wedge \neg Ay))$ Premise

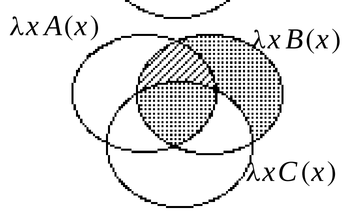
(4) $\exists x(Ax \vee Bx)$ Premise

(5) 

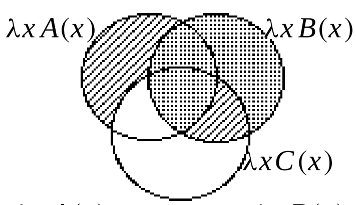
by Rule of Setup

(6) 

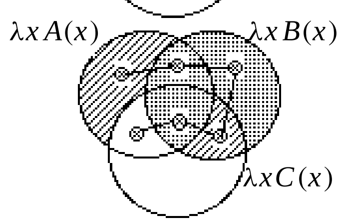
by ∀-Apply from (1) and (5)

(7) 

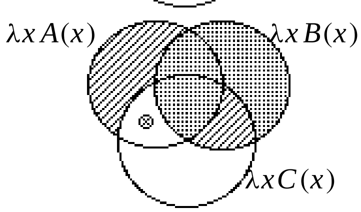
by ∀-Apply from (2) and (6)

(8) 

by ∀-Apply from (3) and (7)

(9) 

by ∃-Apply from (4) and (8)

(10) 

by Erasure of Links from (9)

(11) $\exists x(Ax \wedge Cx \wedge \neg Bx)$ by ∃-Observe from (10)

Figure 6: A heterogeneous proof in the system.

7.3 First-Order Rules The focus of the present paper is on the diagrammatic and heterogenous aspects of the system. Therefore, rather than providing a particular set of first-order rules and axioms, I'll simply assume that we have an axiomatization which is sound and complete with respect to $\models_{\textit{first-order}}$.

Having now stated all the rules and axioms of the system, Figure 6 shows a proof in the system of the last line from the first four lines. The two styles of shading are used to highlight the region of focus and the rectangles are omitted.

8 Soundness

Lemma 8.1 *If Γ is a set of sentences and φ is a sentence, then $\Gamma \models \varphi$ if and only if $\Gamma \models_{\textit{first-order}} \varphi$.*

Proof. This follows from the fact that every first-order model can be extended to a model, and that truth of a sentence in a model depends only on the first-order fragment of the model.

Theorem 8.2 Soundness Theorem *Let Δ be a set of wfos and δ be a wfo. Then, if $\Delta \vdash \delta$ then $\Delta \models \delta$.*

Proof. By induction using the fact that each of the rules is valid. By a standard argument, all the first-order rules are valid. For the remaining rules:

Unification: Let D'' result from D and D' by this rule, and let D and D' be true in (U, I). Then every region of D'' has a counterpart in either D or D', and any shading or sequence occurring in a region of D'' occurs also in its counterpart in either D or D'. Therefore, since I assigns each of these regions of D or D' the empty set, a non-empty set, or a set containing a particular object according to whether it is shaded, has an X- sequence, or a constant-sequence, by the counterpart lemma we have that D'' is true in (U, I).

$\forall$-Apply: Let $(U, I) \models \forall x\varphi$, $(U, I) \models D$, $r \cong \lambda x\varphi$, and D' be obtained from D by the shading of some subregion of $\overline{r}$ or the addition of a sequence to r. By the interpretation lemma, $I(r) = I(\lambda x\varphi)$. Further, since $(U, I) \models \forall x\varphi$, we have that $(U, I) \models \varphi[v]$ for every assignment function v in U. Thus, by the definition of extended first-order model, $I(\lambda x\varphi) = U$. So $I(r) = U$. By the definition of model, then, $I(\overline{r})$ is empty. By the counterpart lemma, $\overline{r}$'s counterpart in D' is also assigned the empty set. Therefore, whatever subregion of $\overline{r}$'s counterpart in D' is shaded in applying the rule, we have that $(U, I) \models D'$. Also, since U is non-empty, $I(r)$ is non-empty and contains the denotation of every constant symbol. Therefore, if a sequence is added to r to get D' we still have, using the counterpart lemma, that $(U, I) \models D$.

$\forall$-Observe: Let $(U, I) \models D$, $r \cong \lambda x\varphi$, and $\overline{r}$ be shaded. By the interpretation lemma, $I(r) = I(\lambda x\varphi)$. Further, $U - I(r) = I(\overline{r})$, which is empty since $(U, I) \models D$. So $I(r) = U$. So $I(\lambda x\varphi) = U$. By definition of extended first-order model, $\{d \in U : (U, I) \models \varphi[v]$ for some assignment function v assigning d to $x\}$ $= U$. By the definition of truth in a model, $(U, I) \models \forall x\varphi$. The remaining cases are verified similarly, concluding the proof.

9 Completeness Completeness of the system is proved by exploiting the fact that the first-order fragment of the rules constitute a complete system with respect to $\models_{\textit{first-order}}$. The representation lemma allows this standard result to be used by showing that every diagram has a linguistic counterpart that is provably and semantically equivalent to it.

Lemma 9.1 Representation Lemma *For every diagram D there is a formula φ such that: (1) φ is provable from D, (2) D is provable from φ , (3) D is a logical consequence of φ , and (4) φ is a logical consequence of D.*

Proof. Let D be a diagram. Let s be the shaded region of D. Let $r_1, ..., r_m$ be the regions of D with X-sequences. Let $t_1, ..., t_n$ be the regions of D with a_i-sequences for $1 \leq i \leq n$ where each a_i is some constant symbol. Let $s \cong \lambda x S(x)$, $r_i \cong \lambda x R_i(x)$ for each $1 \leq i \leq m$, $t_i \cong \lambda x T_i(x)$ for each $1 \leq i \leq n$. Then the desired formula φ is:

$$\neg\exists x S(x) \wedge \exists x R_1(x) \wedge .... \wedge \exists x R_m(x) \wedge T_1(a_1) \wedge .... \wedge T_n(a_n).$$

By the soundness theorem, to prove the representation lemma it suffices to prove (1) and (2). To prove (1), observe that since s is shaded and $s \cong \lambda x S(x)$, it follows that $\overline{s} \cong \lambda x \neg S(x)$. Thus, we get from D by $\forall$-Observe $\forall x \neg S(x)$. Then, using first-order rules we can get $\neg\exists x S(x)$. Since each r_i has an X-sequence, and since $r_i \cong \lambda x R_i(x)$ for each $1 \leq i \leq m$, we can infer $\exists x R_i(x)$ for each $1 \leq i \leq m$ using the rule $\exists$-Observe. Similarly, we get $T_i(a_i)$ for each $1 \leq i \leq n$ from D using the rule Constant Observe. Then, first-order rules yield the sentence φ, completing the derivation of φ from D. To prove (2), use the setup rule to get a diagram with the same tagged curves as D. The first-order rules give $\forall x \neg S(x)$, which allows s to be shaded by $\forall$-Apply. Furthermore, each $\exists x R_i(x)$ and $\exists$-Apply allows an X-sequence to be put into each region r_i, $1 \leq i \leq n$. Finally, each $T_i(a_i)$ and the rule Constant-apply allows an a_i-sequence to be put in each region t_i, $1 \leq i \leq m$. This completes the proof.

Lemma 9.2 Equivalence of $\models$ and $\vdash$ for sentences *Let Γ be a set of sentences and φ be a sentence. Then $\Gamma \models \varphi$ if and only if $\Gamma \vdash \varphi$.*

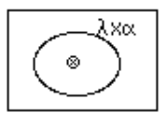
Proof. Suppose $\Gamma \models \varphi$. Then $\Gamma \models_{\textit{first-order}} \varphi$. So by the fact that $\models_{\textit{first-order}}$ and $\vdash_{\textit{first-order}}$ coincide we have that $\Gamma \vdash_{\textit{first-order}} \varphi$. Therefore $\Gamma \vdash_{\textit{first-order}} \varphi$ and so $\Gamma \vdash_{\textit{first-order}} \varphi$. The other direction follows from soundness.

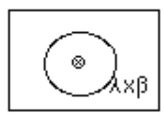
Theorem 9.3 Completeness Theorem *Let Δ be a set of closed well-formed representations and δ be a closed well-formed representation. Then $\Delta \models \delta$ if and only if $\Delta \vdash \delta$.*

Proof: For any closed wfr γ or set of wfrs Γ, let $[\gamma]$ and $[\Gamma]$ be the sentence and set of sentences, respectively, that result from replacing each diagram occurring in $\Gamma \cup \{\gamma\}$ with its linguistic equivalent in accordance with the representation lemma. Suppose $\Delta \models \delta$. By the representation lemma $[\Delta] \models \alpha$ for every $\alpha \in \Delta$, and also $\delta \models [\delta]$. So $[\Delta] \models [\delta]$. By the previous lemma, it follows that $[\Delta] \vdash [\delta]$. Let $[\Delta_0]$ be the finite subset used. Then $[\Delta_0] \vdash [d]$. By the representation lemma $\Delta_0 \vdash \alpha$ for every $\alpha \in [\Delta_0]$, and also $[\delta] \vdash \delta$. Hence $\Delta_0 \vdash \delta$. But then since $\Delta \vdash \alpha$ for every $\alpha \in \Delta_0$, we have $\Delta \vdash \delta$, as desired.

10 Isolating the Diagrammatic Fragment of the System The above proof of completeness shows that the first-order rules along with the heterogeneous rules and the setup rule form a complete system. In other words, it shows that none of the diagrammatic rules save setup are needed for completeness. Nonetheless, the diagrammatic rules often allow one to give more natural proofs between diagrams. For example, to eliminate a link of a sequence falling in a shaded region without the use of the corresponding diagrammatic rule would require transferring information of the first diagram into sentential form, carrying out a first order proof from these sentences to the linguistic equivalent of the second diagram, and finally applying this information back into diagrammatic form. A result is therefore needed showing that the purely diagrammatic rules of the system are complete with respect to some natural notion of "diagrammatic consequence," $\models_{Diag}$, holding among diagrams.

It is not sufficient to define $D \models_{Diag} D'$ as holding whenever $D \models D'$ and both D and D' are diagrams. For let α and β be any two sentences such that β is a logical consequence of α. Then, trivially, the right-hand diagram below is a logical consequence of the left- hand diagram. But clearly this fact does not hold in virtue of the two diagrams' "diagrammatic" features. Rather, it holds only because of the logical structure of the two set terms. A case of logical consequence between sentences has merely been superficially coded up as a question about diagrams. So the desired notion $\models_{Diag}$ will need to ignore such spurious examples.





The desired notion is obtained by treating the labels as atomic in the following sense. Redefine "model" to allow basic regions to be interpreted as any set, provided only that any two basic regions sharing the same label are interpreted to represent the same set. So the only feature of labels now that is semantically relevant is whether or not the same label tags two curves. No analysis beyond mere identity is needed. Then, let models be defined as before on non-basic regions, treating basic regions as just described. In other words:

Definition 10.1 A pair (U, I) is a "diagrammatic model" if and only if U is a set and I a function from regions into the powerset of U satisfying:

1. Any two basic regions enclosed by closed curves having the same label are assigned the same subset of U, and
2. If r and s are regions of a diagram D, then: $I(r \cup s) = I(r) \cup I(s)$, $I(r \cap s) = I(r) \cap I(s)$, $I(r - s) = I(r) - I(s)$, and $I(\overline{r}) = U - I(r)$.

Let $\models_{Diag}$ be the resulting notion of logical consequence between diagrams arising from this new definition of model. Let $\vdash_{Diag}$ be the syntactic notion holding between a set Δ of diagrams and a diagram D if and only if there is a proof of D from Δ using only the diagrammatic rules of the system with the exception of the Rule of Non-Emptiness (which is just a concession to the non-emptiness assumption of first-order logic).

In [3], Shin proved that for any finite set $\Delta \cup \{D\}$ of diagrams, $\Delta \models_{Diag} D$ if and only if $\Delta \vdash_{Diag} D$. Hammer and Danner [2] extended this result to infinite sets, proving that for any set $\Delta \cup \{D\}$ of diagrams, $\Delta \models_{Diag} D$ if and only if $\Delta \vdash_{Diag} D$. This theorem provides the desired result, showing that the diagrammatic rules of the system are complete with respect to the "diagrammatic" notion of consequence just defined. So the system is "diagrammatically complete" in the sense that whenever a diagram can be seen to follow from a set of diagrams without analyzing labels beyond being able to recognize the identity of two labels, it can be proved from that set using only diagrammatic rules of the system.

Acknowledgment I would like to thank Sun-Joo Shin, Jon Barwise, and an anonymous referee for helpful comments on this paper.

REFERENCES

[1] Boolos, George., "A Curious Inference," *The Journal of Philosophical Logic*, vol. 16 (1987), pp. 1-12.

[2] Hammer, Eric., and Norman Danner, "Towards a Model Theory of Venn Diagrams," Unpublished Paper.

[3] Shin, Sun-Joo., *The Logical Status of Diagrams*, Cambridge University Press, Cambridge, forthcoming.

[4] Shin, Sun-Joo., "A Situation-Theoretic Account of Valid Reasoning with Venn Diagrams," pp. 581–605 in *Situation Theory and its Applications II*, CSLI, 1991.

Department of Philosophy
Indiana University
Bloomington, Indiana 47405
e-mail:ehammer@ucs.indiana.edu