A Diagrammatic Proof Search Procedure as Part of a Formal Approach to Problem Solving*

Diderik Batens
Centre for Logic and Philosophy of Science
Universiteit Gent, Belgium
Diderik.Batens@ugent.be

Abstract

This paper aims at describing a goal-directed and diagrammatic method for proof search. The method (and one of the logics obtained by it) is particularly interesting in the context of formal problem solving. A typical property is that it consists of attempts to justify so-called bottom boxes by means of premise elements (diagrammatic elements obtained from premises) and logical elements. Premises are not preprocessed, whence most premises lead to a variety of premise elements.

The method is simple and insightful in three respects: (i) diagrams are constructed by drawing the goal node and superimposing the top node of a new diagrammatic element on a bottom box of an element that occurs in the diagram; (ii) diagrammatic elements are built up from binary and ternary relations that connect nodes (comprising one or two boxes) to boxes (entities containing a single formula); (iii) diagrammatic elements are obtained in view of existing bottom boxes by a unified approach. At the propositional level, the method is an algorithm for derivability (but leaves choices to the user). Extended to the predicative level, it provides a criterion for derivability and one for non-derivability.

The method is demonstrably more efficient than tableau methods and has certain advantages over linear methods and certain other goal-directed methods. Apart from making certain properties of search paths more visible, the method also led to a simplification of the metatheoretic proofs.

1 Aim of this paper

In [3] and [2], I spelled out the basics of a formal approach to problem solving. An important part of it consists of a proof search method that is largely pushed into the proofs.[1] Some first results were meanwhile published—see for example [6], [1] and [4].

---

*Research for this paper was supported by subventions from Ghent University and from the Fund for Scientific Research – Flanders, and indirectly by the Flemish Minister responsible for Science and Technology (contract BIL01/80). I am indebted to Joke Meheus for useful comments on a former draft.

[1] One objection against a strict distinction between rules of inference ("definitory rules") and heuristic rules ("the strategy")—see for example [8]—is that heuristic reasoning is itself a form of reasoning.

The approach followed in those papers is prooflike, and there are a number of arguments in favour of this. However, both for finding a correct formulation and for finding the proofs of the metatheorems, it turned out extremely helpful to think about the proof search in terms of a tree-like structure. Moreover, I often had the impression that much insight could be gained by explicitly spelling out a diagrammatic version of the proof search procedure. The present paper contains a (semi-formal) formulation of this version as well as some first results that derive from it.

2 Two Examples

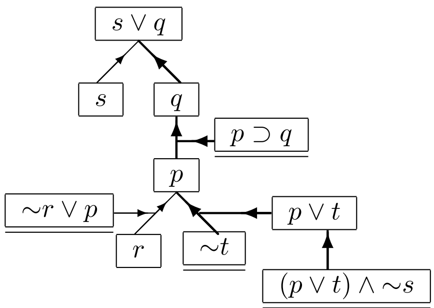
Suppose that one tries to derive $s \vee q$ from the premises

$$\sim t,\ p \supset q,\ \sim r \vee p,\ p \supset \sim s,\ (p \vee t) \wedge \sim s\,.$$

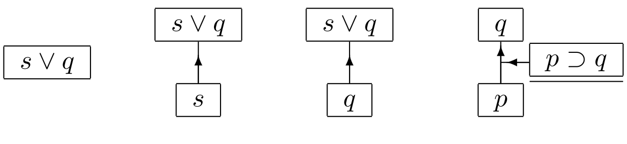
The search process may have been the following. Clearly, $s \vee q$ cannot be obtained directly from any premise. So one tries to derive either s or q. While s cannot be obtained from a premise, the second premise tells us that q can be obtained if p can be obtained, whence one looks for p. In view of the third premise, p can be obtained if r can be obtained. But r clearly cannot be obtained from any premise. So one looks for another way to obtain p. The fifth premise entails $p \vee t$; in view of this, p can be obtained if $\sim t$ can be obtained. And $\sim t$ is the first premise. So one is home.

The previous paragraph *describes* the search process, but does not make it fully explicit. Thus the claim that $s \vee q$ cannot be obtained directly from any premise implicitly states that all five premises have been considered. So, although the fourth premise is never explicitly mentioned in the description of the search process, it was considered repeatedly. However, the fourth premise failed to be useful at any point, and hence did not lead to any search path. The third premise led to a search path, which however was unsuccessful.

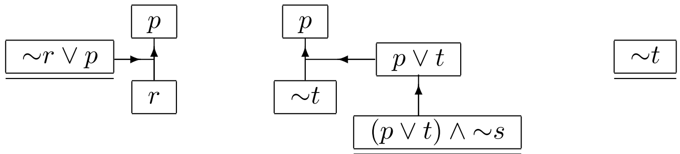
In view of what was said in Section 1, I now present the proof search in a diagrammatic way.



The diagram can best be seen as composed of seven (overlapping) diagrammatic elements. The top half of the diagram is constructed from the following elements:

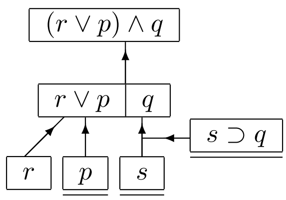


The leftmost element is the *goal element*. The diagram is started by drawing it. The next two elements are analysing elements or *logical elements*, expressing respectively that $s \vee q$ is obtained from s and that $s \vee q$ is obtained from q. The fourth element is a *premise element*—the premise occurs in the underlined node. The element represents that $p \supset q$ is a premise and that it justifies the transition from p to q. Here are the other elements from which the diagram is constructed:



All of these are premise elements—obtained from premises in a way taht is spelled out in Section 3.

Sometimes there is a slight complication, as is illustrated by the diagram for $s, p, s \supset q \vdash (r \vee p) \wedge q$ (the diagram obtained in trying to derive $(r \vee p) \wedge q$ from the premise set $\{s, p, s \supset q\}$).



No premise element has $(r \vee p) \wedge q$ as its top node (see Section 3), and hence the diagram cannot be extended downward from the goal node by means of such an element. So the goal node has to be analysed by means of a logical element. This analysis gives us a two-box node: $(r \vee p) \wedge q$ is obtained if both $r \vee p$ and q are obtained. Remark that two diagrammatic elements are ‘attached’ to the box containing $r \vee p$, whereas only one element is ‘attached’ to the box containing q. In general, all nodes contain either one or two *boxes*, and every box contains a single formula.

We have seen that there are three kinds of diagrammatic elements: the goal element, logical elements, and premise elements (a premise node or a set of connected nodes obtained from a premise). It is useful to distinguish within the elements *top nodes* (nodes from which no arrow departs) and *bottom nodes* (nodes in which no arrow arrives). In single node elements, the top node and bottom node coincide. For the other elements, there is exactly one top node and there are one or more bottom nodes.

A diagram is constructed starting from the top, by drawing the goal node, and is extended downward by superimposing the top node of a new element on a *box* of a bottom node that already occurs. It is important to stress this: a diagram is extended downward from a bottom box in an attempt to justify the formula that occurs in the box. If a bottom node comprises two boxes, both formulas have to be justified.

There are two kinds of *arrows* in the diagram. A simple arrow, for example the one from s to $s \vee q$, is a *logical arrow*. It corresponds to $s \vdash s \vee q$ (or, where Γ denotes the premise set, to “if $\Gamma \vdash s$, then $\Gamma \vdash s \vee q$”). A combined arrow (an arrow at which another arrow arrives) is a *contingent arrow*—the combined arrow represents a ternary relation. Thus the contingent arrow between $p \supset q$, p and q expresses that the transition from the minor p to the conclusion q is warranted by the major $p \supset q$.[2] It corresponds to $p \supset q, p \vdash q$ (or to "if $\Gamma \vdash p \supset q$, then if $\Gamma \vdash p$, then $\Gamma \vdash q$"). Incidentally, in the diagrams presented in this paper, there is at most one arrow to an arrow, Contingent arrows occur only in premise elements, and the arrow of a logical element is always a logical arrow.

Several paths may be distinguished on a diagram. For the time being, a path can be seen as a chain of diagrammatic elements. A path is successful if (again, for the time being) all its bottom boxes (boxes in which no arrow arrives) are premises, which is seen on the diagrams by the fact that they are underlined—more precise definitions follow in Section 3.

Once a successful path on the diagram is identified, there is an algorithm for transforming it to, for example, a Fitch-style proof—see [9] for some first results. While the search three was obtained by moving down from the goal node (see the paragraph on the search process), the proof is stepwise obtained by moving up along the only successful search path in the first diagram.

1	$(p \vee t) \wedge {\sim}s$	Premise
2	$p \vee t$	1; Simplification
3	${\sim}t$	Premise
4	p	2, 3; Disjunctive Syllogism
5	$p \supset q$	Premise
6	q	4, 5; Modus Ponens
7	$s \vee q$	6; Addition

The following proof is derived from the successful path of the second diagram.

1	p	Premise
2	$r \vee p$	1; Addition
3	s	Premise
4	$s \supset q$	Premise
5	q	3, 4; Modus Ponens
6	$(r \vee p) \wedge q$	2, 5; Adjunction

If one compares the first proof to the heuristic reasoning, the most striking feature is that several search paths were tried but do not occur in the proof. Proofs are supposed to demonstrate that a conclusion can be obtained from the premises, and they are supposed to do so in an elegant way. This is the reason why the third premise (from the list in the first paragraph of this section) is not even introduced in the proof. More importantly, the proof contains no trace of the unsuccessful search for r that was induced by the third premise.

3 Diagrammatic Elements, Paths, Successful Paths, and Descendants

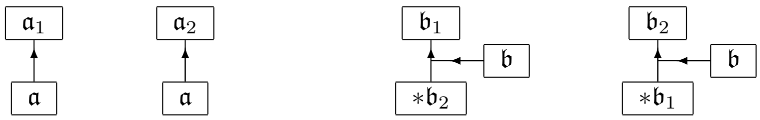
In order to systematically and concisely describe the way in which logical elements and premise elements are obtained, I distinguish between $\mathfrak{a}$-formulas and

[2] Nothing prevents that $s \supset (s \vee q)$ is a premise, justifying the transition from s to $s \vee q$.

$\mathfrak{b}$-formulas (varying on a theme from [10]). Let $*A$ denote the 'complement' of A, viz. B if A is $\sim B$ and $\sim A$ otherwise. To each 'complex' formula two other formulas are assigned according to the following table:

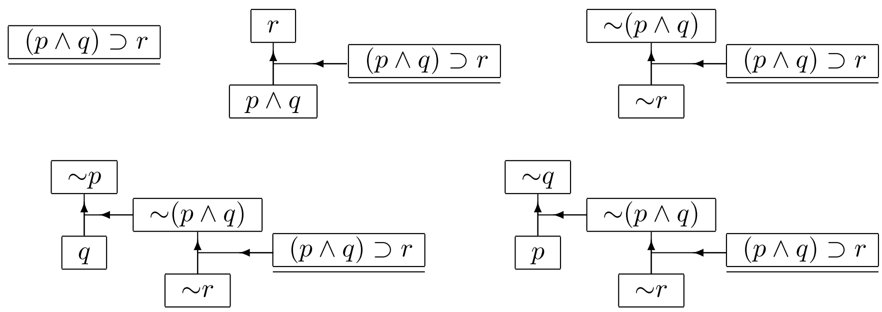
$\mathfrak{a}$	$\mathfrak{a}_1$	$\mathfrak{a}_2$		$\mathfrak{b}$	$\mathfrak{b}_1$	$\mathfrak{b}_2$
$A \wedge B$	A	B		$\sim(A \wedge B)$	$*A$	$*B$
$A \equiv B$	$A \supset B$	$B \supset A$		$\sim(A \equiv B)$	$\sim(A \supset B)$	$\sim(B \supset A)$
$\sim(A \vee B)$	$*A$	$*B$		$A \vee B$	A	B
$\sim(A \supset B)$	A	$*B$		$A \supset B$	$*A$	B
$\sim\sim A$	A	A				

Most premises generate several diagrammatic elements. First of all, each premise gives us a one-node premise element. Next, a premise element may be extended by a so-called upward move to its top node. The *upward moves* for $\mathfrak{a}$-formulas and $\mathfrak{b}$-formulas are respectively as follows:

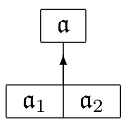


So both $\mathfrak{a}$-formulas and $\mathfrak{b}$-formulas may be extended in two ways, depending on whether one wants to obtain the first or the second associated formula. An important difference is that an $\mathfrak{a}$-formula justifies both its associated formulas, whereas a $\mathfrak{b}$-formula justifies the transition from the complement of one of its associated formulas to the other associated formula. Extending $\mathfrak{a}$-formulas upward leads to logical arrows. Extending $\mathfrak{b}$-formulas upward leads to contingent arrows.

If the top node of the thus obtained element is itself an $\mathfrak{a}$-formula or a $\mathfrak{b}$-formula, two new diagrammatic elements may be obtained by a further upward move, and so on. It is instructive to list the diagrammatic elements that are obtained from a premise, for example $(p \wedge q) \supset r$, which delivers the following five elements:

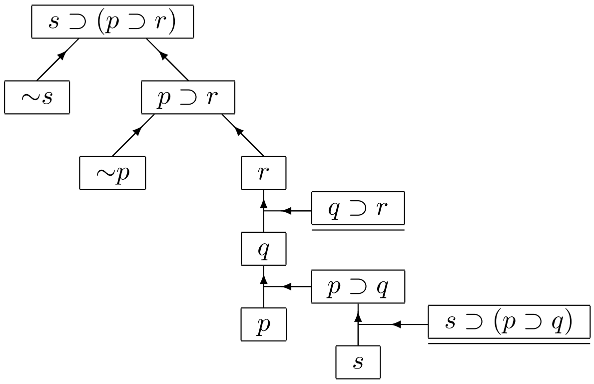


Let us now turn to the way in which logical elements are obtained. We have seen that $\mathfrak{a}$-formulas are justified by their associated formulas *taken together*. *Downward move for* $\mathfrak{a}$*-formulas*:



For $\mathfrak{b}$-formulas, the following downward extensions are certainly correct:

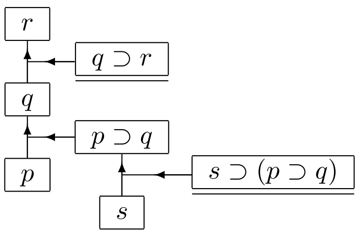
However, this approach does not provide a complete search system for Classical Logic (henceforth **CL**). One of the reasons for this is that Excluded Middle is absent from it. This can be seen from the diagram for $q \vdash p \vee \sim p$, but it is instructive to consider an example that illustrates the problem in a more general way, viz. the diagram for $s \supset (p \supset q), q \supset r \vdash s \supset (p \supset r)$.



On the present criteria, there is no successful path (no path of which all nodes are justified). Yet the conclusion is derivable from the premises by **CL**. And indeed, this can be seen from the diagram as follows. Either p or $\sim p$ is true, and either s or $\sim s$ is true. If $\sim s$ is true, $s \supset (p \supset r)$ is true by the leftmost path of the diagram. If $\sim p$ is true, $s \supset (p \supset r)$ is true by the middle path of the diagram. And if both p and s are true, then $s \supset (p \supset r)$ is true by the rightmost path of the diagram.

So, it seems that we need a supplementary criterion for justified nodes or paths. The criterion that is implicitly used in the previous paragraph—let us call it the EM (excluded middle) criterion—seems to offer just the required way out. And yet, there is a more attractive approach.

Consider the following fragment (of the rightmost path) of the diagram:

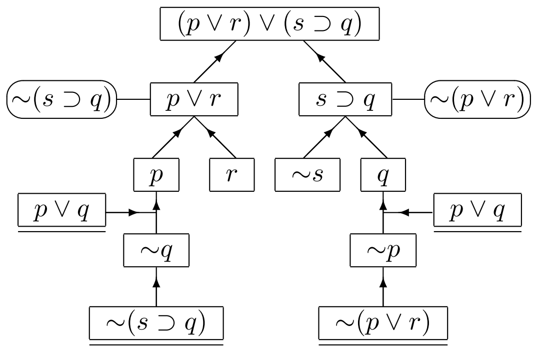


This fragment shows that, if s and p were premises, then r would be derivable: $s \supset (p \supset q), q \supset r, s, p \vdash r$. By applying the deduction theorem twice to this, one obtains $s \supset (p \supset q), q \supset r \vdash s \supset (p \supset r)$ as desired. This criterion is more appealing and more intuitive than the EM criterion, but it sometimes fails, as may be seen from the diagram for $p \vee q \vdash (p \vee r) \vee (s \supset q)$—it is part of the next diagram. However, the criterion shows us the way to a better approach.

When precisely does one want to apply a supplementary criterion? This happens in cases in which neither $\mathfrak{b}_1$ nor $\mathfrak{b}_2$ is derivable from the premises, whereas $\mathfrak{b}$ is. Thus neither $\sim s$ nor $p \supset r$ is derivable from $\{s \supset (p \supset q), q \supset r\}$, but $s \supset (p \supset r)$ is. However, we know that $\Gamma \vdash \mathfrak{b}$ holds true iff $\Gamma \cup \{*\mathfrak{b}_2\} \vdash \mathfrak{b}_1$, and also iff $\Gamma \cup \{*\mathfrak{b}_1\} \vdash \mathfrak{b}_2$ holds true. So the introduction of a new criterion for justified nodes can be avoided if one allows that $*\mathfrak{b}_2$ is used as a supplementary premise on the paths on which one tries to justify $\mathfrak{b}_1$, and that $*\mathfrak{b}_1$ is used as a supplementary premise on the paths on which one tries to justify $\mathfrak{b}_2$. This leads to the following *downward moves for* $\mathfrak{b}$*-formulas*:

in which a formula in an 'oval', attached to a node, indicates that this formula may be used as a supplementary premise on the paths to which this node belongs—it will only be used 'below' this node, whence no ambiguity arises.

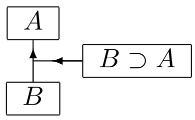
Let us see what becomes of the diagram for $p \vee q \vdash (p \vee r) \vee (s \supset q)$ in view of the downward move for $\mathfrak{b}$-formulas—I do not draw the additional premises on the third level from the top as they are useless anyway.



Both outer paths (the first and the fourth one) are successful. Of course there is no need to even start the second path once one found that the first path was successful, but I drew all possible paths in order to show that the previous criterion (the one without the new premises) does not work.[3]

This ends the description of the way in which logical elements and premise elements are obtained. It is important to realize that these elements are restricted in such a way that they lead only to the analysis of targets (formulas one tries to justify) and to the analysis of premises. Thus, if A is the formula of a bottom box of a diagrammatic element, no move enables one to introduce, for example, the following 'logical' element:

[3]A successful path on which a supplementary premise is used, is most naturally turned into a Fitch-style proof by starting a subproof with the supplementary premise as its hypothesis, applying Conditional Proof once the desired formula is obtained, and next transforming the result as required.

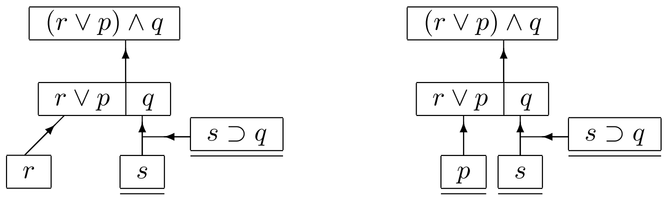


It obviously holds that, where Γ denotes the premise set, "if $\Gamma \vdash B \supset A$, then, if $\Gamma \vdash B$, then $\Gamma \vdash A$". However, introducing boxes containing B and $B \supset A$ in order to justify A is an arbitrary and inefficient move, not a goal-directed move. For this reason, the aforementioned restriction on logical elements and premise elements is essential. In Section 7, I shall mention a further (optional) restriction on premise elements, which is directed towards the efficiency of the search process.

Before turning to to the definition of a path on an diagram and of a successful path on a diagram, I need to specify the relations represented by the arrows. I shall do this in an informal way, leaving the obvious precise mathematical description to the reader.

A logical arrow connects a node to a box; a contingent arrow connects two nodes to a box. Apart from what is obvious, the essential further convention is that, while many arrows may end up in the same box, it is never the case that two arrows (whether logical or contingent) depart from a node. Of course, nothing prevents one to use the same diagrammatic element at different points in the same diagram.

A *path* on a diagram is a smallest set of connected diagrammatic elements defined by: (i) the goal node belongs to every path, (ii) if a box is the top node of n different diagrammatic elements, then each of these elements belongs to a different path (leading from the element to the goal node). There are three paths in the first diagram of Section 2. It may be less obvious that there are two paths in the second diagram of that section, viz.:



The node that comprises two boxes belongs to both paths. On each path, each of these boxes is the top node of a diagrammatic element.

A path is *justified* iff (if and only if) all its nodes are justified. A node is justified iff all its boxes are justified. That a box is justified (in view of the premises) is recursively defined by the following clauses:

1. A node is justified if all its boxes are justified.

2. Premise boxes (underlined boxes) are justified.

3. A box that is reached by a logical arrow from a justified node or by a contingent arrow from two justified nodes is justified.

There is one justified path in each diagram of Section 2—in the first diagram, the path is marked by bold arrows, in the second diagram the justified path is the rightmost one in the figure that displays them separately (in this section).

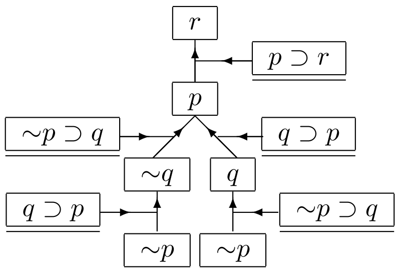
A *bottom box of a diagram* is a box in which no arrow arrives; a *bottom box of a path* is a bottom box of the diagram that belongs to the path. These should not be confused with the bottom boxes of a diagrammatic element.

In the sequel, I shall also need the notion of a descendant of a box. That a box is a *descendant* of another box is recursively defined by (i) box b is a descendant of box b' if there is a logical arrow from the node containing b to the box b', (ii) box b is a descendant of box b' if there is a contingent arrow from the node containing b (and from another node) to the box b', and (iii) if box b is a descendant of box b' and box b' is a descendant of box b'', then box b is a descendant of box b''.

If two boxes occur on the same path, it is possible that neither is a descendant of the other. This is due to the occurrence of contingent arrows and to the downward move for $\mathfrak{a}$ formulas. Thus, in the first diagram of Section 2, the box containing $(p \vee t) \wedge {\sim} s$ is not a descendant of the box containing ${\sim} t$, nor is the latter box a descendant of the former. In the second diagram of Section 2, the descendant relation does not obtain in either direction between, for example, the box containing p and the box containing s.

4 Reasoning *Ex Absurdo*

The method described up to now still characterizes a logic that is weaker than Classical Logic (henceforth **CL**). One of the reasons for this can be seen from the diagram for ${\sim} p \supset q, q \supset p, p \supset r \vdash r$.



On the present criteria, there is no successful path in this diagram, whereas r is **CL**-derivable from $\{{\sim} p \supset q, q \supset p, p \supset r\}$. Indeed, this may be seen from the diagram. Actually, both paths show that ${\sim} p$ is sufficient to justify p. But if $\Gamma \cup \{{\sim} p\} \vdash_{\mathbf{CL}} p$, then $\Gamma \vdash_{\mathbf{CL}} p$.

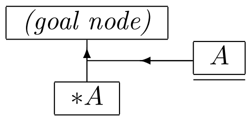
In order to build this in into the search method, the definition of a justified box needs to be extended with: *if box* b *is a descendant of box* b'*, and box* b *contains the complement of the formula of box* b'*, then box* b *is justified.*

Let us return to the last diagram. In view of the change, the two bottom boxes containing ${\sim} p$ are justified. That the box containing ${\sim} p$ is justified has the effect that the box containing p is also justified. In other diagrams, there may be unjustified boxes containing other formulas; an example is the diagram for ${\sim} p \supset q, q \supset p, s \supset (p \supset r) \vdash r$, which, even on the modified definition, contains no justified path, as required.

5 Where Went Explosion?

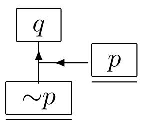
The Explosion rule, also known as *ex falso quodlibet*, states that every formula can be derived from an inconsistency ($A, {\sim}A/B$). This rule warrants that, in **CL** and other systems that share the rule, $p, {\sim}p \vdash q$. In its present state, the diagrammatic method does not validate Explosion. Thus the proof search diagram for $p, {\sim}p \vdash q$ consists only of the goal node, containing q; no diagrammatic element obtained from the premises has q as its top node.

It is not difficult to add a further (and final) move to handle Explosion. The idea is that, if A is a premise ($A \in \Gamma$), and $\Gamma \vdash {\sim}A$, then anything can be derived from the premise set ($\Gamma \vdash B$ for every B). This is expressed in diagrammatic terms by the following move, which may be seen as an upward move generated by a premise $A \in \Gamma$:



As the display suggests, I restrict this move by requiring that the top node of the element is the goal node. Although this element could in principle be attached to any box, the only sensible way of proceeding consists in attaching it to the goal node.

The procedure then looks as follows: first one tries the method described in the previous sections. If this fails, one applies the Explosion move to the first premise, next to the second premise, and so on. A very simple example is the diagram for $p, {\sim}p \vdash q$:



This completes the description of the proof search diagrams. The following theorems can be proved about the method.

Theorem 1 *If* Γ *is finite, then so is the diagram for* $\Gamma \vdash A$.

Theorem 2 $\Gamma \vdash A$ *iff the diagram for* $\Gamma \vdash A$ *contains a successful path.*

If the procedure is upgraded from propositional **CL** to full **CL**, Theorem 1 obviously fails (full **CL** is undecidable). Still, Theorem 2 may be adapted to the following: if the diagram is finite and contains a successful path, then $\Gamma \vdash A$; if the diagram is finite and contains no successful path, then $\Gamma \nvdash A$.

6 The Logic CL$^-$

The Explosion move is isolated from the others. What happens if one does not add it to the procedure?

Not to apply the Explosion move is often sensible. Indeed, in many cases one does not want Explosion. Thus, if one is trying to derive a prediction from a theory together with a set of data, one is not interested in any prediction that is derivable by Explosion—if A is so derivable, so is $\sim A$. No one will act on A in this case, and A will not be considered as part of the predictive power of the theory—similarly for the explanatory power of the theory.

Let $\mathbf{pCL}^-$ be the diagrammatic method without the move handling Explosion, and let $\Gamma \vdash_{\mathbf{pCL}^-} A$ iff the method without the Explosion move leads to a diagram for $\Gamma \vdash A$ in which at least one path is successful. One can then define the logic $\mathbf{CL}^-$ as follows:

Definition 1 $\Gamma \vdash_{\mathbf{CL}^-} A$ *iff* $\Gamma \vdash_{\mathbf{pCL}^-} A$.

The following theorems can be proved—see a forthcoming paper.

Theorem 3 *If* Γ *is finite, then so is the diagram for* $\Gamma \vdash_{\mathbf{CL}^-} A$.

Incidentally, means to handle infinite premise sets are described in [6].

Theorem 4 *If* Γ *is consistent, then, for all* A, $\Gamma \vdash_{\mathbf{CL}^-} A$ *iff* $\Gamma \vdash_{\mathbf{CL}} A$.

Theorem 5 *If* Γ *is inconsistent, then there is an* A *such that* $\Gamma \vdash_{\mathbf{CL}^-} A \wedge \sim A$.

In other words, $\mathbf{CL}^-$ leads to exactly the same consequences as **CL** if the premise set is consistent (this is the intended domain of application of **CL**), whereas every inconsistent premise set has an explicit contradiction as a $\mathbf{CL}^-$-consequence, as desired.

As was shown in [5], $\mathbf{CL}^-$ is sound and complete with respect to a semantics. In that paper some further properties of $\mathbf{CL}^-$ are studied and more motivation for the system is presented.

For years, logicians have claimed that Explosion cannot be isolated in **CL**, because if one wants to avoid Explosion, one needs to give up, for example, either Disjunctive Syllogism or Addition (as well as one of many other plausible inference rules). The procedural approach shows that these claims are wrong. If a logic is defined by a procedure—roughly by the results of a search process—it is possible to isolate Explosion.

7 Devising the Elements One Needs

If the premise set is large, it seems a drawback of the procedure that all diagrammatic elements of all premises have to be prepared before the construction of the diagram can start. Actually, this can easily be avoided, viz. as follows. One starts the diagram with the goal node as before. Next, one considers a non-justified box of a bottom node. Let the box contain the formula A. One then considers a premise B, starting with the first one, and checks whether A is a positive part of B, $\mathrm{pp}(A, B)$, which is recursively defined as follows:

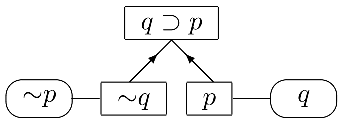
1. $\mathrm{pp}(A, A)$.
2. $\mathrm{pp}(A, \mathfrak{a})$ if $\mathrm{pp}(A, \mathfrak{a}_1)$ or $\mathrm{pp}(A, \mathfrak{a}_2)$.
3. $\mathrm{pp}(A, \mathfrak{b})$ if $\mathrm{pp}(A, \mathfrak{b}_1)$ or $\mathrm{pp}(A, \mathfrak{b}_2)$.
4. If $\mathrm{pp}(A, B)$ and $\mathrm{pp}(B, C)$, then $\mathrm{pp}(A, C)$.

If A, the contents of the box one tries to justify, is a positive part of the premise B, one applies upward moves until the single box of the top node of the diagrammatic element contains A. There is a simple algorithm for obtaining this diagrammatic element, which is then used to extend the diagram downward. If A is not a positive part of any premise, a downward move is applied to A—see Section 3.

8 A Matter of Elegance

The application of the downward move for $\mathfrak{b}$-formulas is somewhat inefficient in that, more often than not, the left move (the one introducing $\mathfrak{b}_1$ with $*\mathfrak{b}_2$ as a supplementary premise) leads to a justified path iff the right move does so. So trying out both paths seems a loss if time.

The trouble is with the "more often than not". Consider the top of the following diagram for $p \vdash q \supset p$.



The right path will obviously be justified (without the supplementary premise), but the left path will not, unless one were to apply the Explosion move to the bottom node containing $\sim q$, which would complicate what was said in Section 5, would not work in $\mathbf{CL}^-$, and would result in an ugly path (and an ugly proof) while nice alternatives are possible. A compromise is obtained by recasting the downward moves for $\mathfrak{b}$-formulas as follows:

The change is justified by the following consideration. Let Γ be the premise set as before and let a $\mathfrak{b}$-formula be analysed according to the modified moves, introducing $\mathfrak{b}_1$ on the left path and $\mathfrak{b}_2$ with the supplementary premise $*\mathfrak{b}_1$ on the right path. If the right path is not successful[4] because $\Gamma \cup \{*\mathfrak{b}_1\}$ is inconsistent, then $\Gamma \vdash \mathfrak{b}_1$, whence the left path is bound to be successful. The choice to add the new premise to the right, rather than to the left, is obviously arbitrary, except that the choice for the right path leads to a nice Fitch-style reconstruction in terms of Conditional Proof.

9 Improving the Efficiency of the Procedure

Not much attention was paid until now to the efficiency of the procedure. My main aim was to spell out a *sensible* procedure that leads to sensible diagrams. But even with respect to sensibility the procedure may be made more (but not completely) deterministic and efficient. The improvement offered in this section concerns the marking of paths.

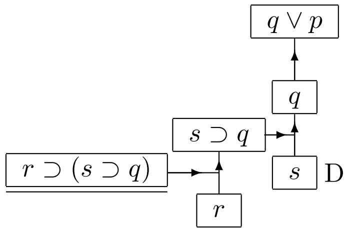
[4]More precisely: if no path to which the node containing $\mathfrak{b}_1$ belongs is successful. Similarly below in the text.

Dead ends At every stage of the diagram, a bottom box containing A will be marked as a *dead end* iff A is not a positive part of any premise and is neither an $\mathfrak{a}$-formula nor a $\mathfrak{b}$-formula (i.e. cannot be analysed by means of a logical element). If this obtains, no logical element and no premise element has a top node containing A.

If a box of a path is marked as a dead end, the path is a *dead end path.* Even if all other bottom nodes on the path can be justified, the marked bottom box will forever remain unjustified. So dead end paths are not continued.

Some care is required here. If a path contains a dead end node n, it is still possible that an higher node of the path may be extended downward, resulting in a new path that does not contain n. The new path may obviously be a justified one. This shows the importance of defining a path in terms of diagrammatic elements. If a node is marked, justifying another bottom node of the same diagrammatic *element* is useless, but justifying a bottom node of an element that is located higher in the path may be useful because it starts a new path.

A simple example of a dead end path occurs if one tries to derive $q \vee p$ from a premise set containing $r \supset (s \supset q)$ and r, but not containing any premise of which either $q \vee p$ or s is a positive part. The diagram might start as follows:



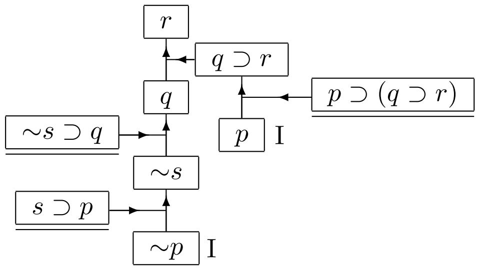
As s is a dead end, one has either to find an alternative justification of q or one has to apply the second downward move for $\mathfrak{b}$-formulas to the goal node.

Inconsistent paths A path is inconsistent iff, for some A, two unjustified boxes that belong to the path contain respectively A and $\sim A$.[5] If this is the case, the lowest box of the two is *marked as inconsistent.* Inconsistent paths are not continued.

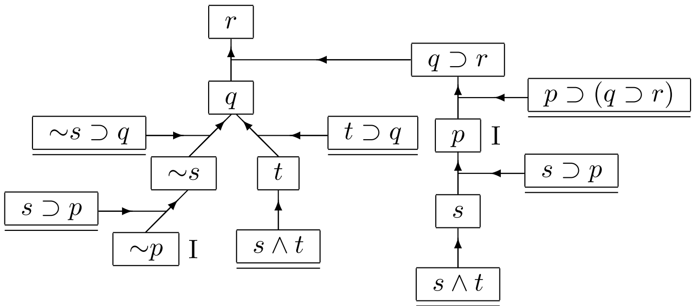
This requires two comments. First, even if the premises are inconsistent, it is still possible that neither A nor $\sim A$ can be justified. So it seems advisable to give up inconsistent paths and, if necessary, to apply Explosion moves. Next, inconsistent paths should *not* be marked in diagrams for the logic $\mathbf{CL}^-$, where they are required for demonstrating the inconsistency of the premise set.

Here is a simple example of an inconsistent search path in a diagram for $p \supset (q \supset r), s \supset p, \sim s \supset q, s \wedge t, t \supset q \vdash r$. The boxes containing respectively p and $\sim p$ are I-marked. This indicates that at least *one* of them cannot belong to a successful path. As $p \supset (q \supset r)$ is the only premise of which r is a positive part, we have to find a path that does not contain $\sim p$. As $s \supset p$ is the only premise of which $\sim s$ is a positive part, we have to find an alternative justification for q, and q is a positive part of $t \supset q$.

[5] The qualification 'unjustified' is required in view of the *Ex Absurdo* justification from Section 4.

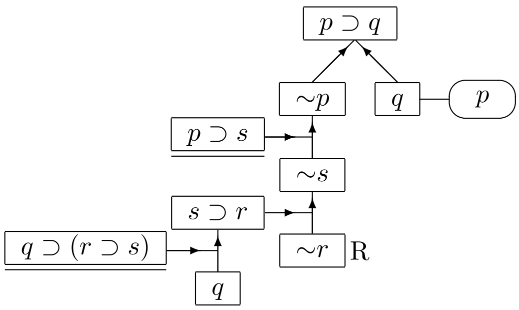


So we have to find an alternative justification for q along this road, which indeed leads to a justified path, viz. the right one.



Redundant paths Where p is a path on a diagram, let $\mathsf{B}(\mathsf{p})$ be the set of unjustified bottom boxes of p. A path p is *marked as redundant at a stage of the diagram* iff, at that stage, there is a path p' such that $\mathsf{B}(\mathsf{p}') \subset \mathsf{B}(\mathsf{p})$ and the premises that are available on path p (including the ones available from downward moves to $\mathfrak{b}$-formulas) are also available on path p'. The underlying idea is obvious: if the members of $\mathsf{B}(\mathsf{p})$ can be justified by the premises, then so can the members of $\mathsf{B}(\mathsf{p}')$, but *not* vice versa. A path is marked as redundant by marking (with a R) an unjustified bottom node from which departs the 'lowest' arrow of the path. In terms of the previous paragraph, the obvious choice is to mark a node in $\mathsf{p} - \mathsf{p}'$.

By way of an example, suppose one is trying to derive $p \supset q$ from a premise set containing $p \supset r$ and $q \supset (r \supset s)$. If both logical moves are tried on the goal node, the top of the diagram may look as follows:

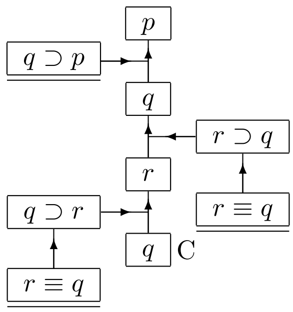


Clearly the right path will be justified if q is justified, whereas the left path is only justified if both q and $\sim r$ are justified. Two roads are open once the left path is R-marked. Either one continues the right path or one looks for an alternative justification for $\sim s$ or for $\sim p$.

Once a path p is marked as redundant at a stage of the construction, there can be no reason to remove the mark at a later stage. The example nicely illustrates this. Even if the justification of the box containing q on the right path would require the justification of $\sim r$, the right path would still be more efficient than the left one in that $\sim r$ has to be justified only once, viz. as part of the justification of q. Indeed, $\sim r$ would have to be justified twice on the left path, once because it is the formula of the R-marked bottom box and once for the justification of the box containing q.

Circular paths A path is *marked as circular* if both box b and box b′ belong to the path, b is a descendant of b′, and b and b′ contain the same formula. The underlying idea is that, if the lower box b can be justified by a chain of diagrammatic elements, then this chain can used to justify the higher box b′, which results in a shorter path. A path is marked as circular by marking (say with a C) the descendant box b.

By way of an example, suppose one is trying to derive p from a premise set containing $q \supset p$ and $r \equiv q$. The diagram might start as follows:



The only node to be justified is the (single box) bottom node containing q. If this node can be justified, the justification may just as well be attached to the higher box containing q. So one either looks for an alternative justification of r, or for an alternative justification of q, or for an alternative justification of p.

10 In Conclusion

The advantages of the diagrammatic approach show if one compares it to the prospective dynamic proofs from [6].

First, the diagrammatic approach has the advantage to clearly identify the paths, and to make the application of the marking definitions extremely transparent. This mainly results from the fact that a prospective element (a formula-plus-condition) may belong to several paths. So defining paths in terms of prospective proofs is tiresome and paths that have to be marked are easily overlooked. A further difficulty is that it is not always clear which prospective elements should be marked in order to indicate that a path has to be marked.

Next, the diagrammatic approach has advantages for the metatheory, for example for the completeness proof—this comes basically to showing that $\Gamma \nvDash A$, viz. that a model of the premises falsifies the conclusion, iff a model falsifies a non-justified bottom node of every path on a completed diagram.

It can be shown that the goal-directedness of the search method makes it more efficient than tableau methods. The method also stands out if compared to the goal directed methods from [7].

In view of what was said in Section 7, the premises need not be preprocessed (for example turned into Horn clauses). An advantage is that if $\Gamma \vdash A$, and Γ' and A' are obtained by substitutions of letters occurring in Γ and A, then (if a decent strategy is followed) the proof for $\Gamma' \vdash A'$ has the same structure as the proof for $\Gamma \vdash A$.

References

[1] Diderik Batens. On a partial decision method for dynamic proofs. In Hendrik Decker, Jørgen Villadsen, and Toshiharu Waragai, editors, *PCL 2002. Paraconsistent Computational Logic*, pages 91–108. (= *Datalogiske Skrifter* vol. 95), 2002. Also available as cs.LO/0207090 at `http://arxiv.org/archive/cs/intro.html`.

[2] Diderik Batens. A formal approach to problem solving. In Claudio Delrieux and Javier Legris, editors, *Computer Modeling of Scientific Reasoning*, pages 15–26. Universidad Nacional del Sur, Bahia Blanca, Argentina, 2003.

[3] Diderik Batens. Notes on problem solving. Technical report, 2003.

[4] Diderik Batens. A procedural criterion for final derivability in inconsistency-adaptive logics. *Journal of Applied Logic*, in print.

[5] Diderik Batens. It might have been Classical Logic. To appear.

[6] Diderik Batens and Dagmar Provijn. Pushing the search paths in the proofs. A study in proof heuristics. *Logique et Analyse*, 173–175:113–134, 2001. Appeared 2003.

[7] Dov M. Gabbay and Nicola Olivetti. *Goal-Directed Proof Theory.* Kluwer, Dordrecht, 2000.

[8] Jaakko Hintikka. Is logic the key to all good reasoning? In *Inquiry as Inquiry: A Logic of Scientific Discovery*, pages 1–24, Dordrecht, 1999. Kluwer.

[9] Dagmar Provijn. How to obtain elegant Fitch-style proofs from goal directed ones. In H. Blockeel and M. Denecker, editors, *Proceedings of the Fourteenth Belgium-Netherlands Conference on Artificial Intelligence*, pages 243–250, Leuven, Belgium, 2002.

[10] Raymond M. Smullyan. *First Order Logic.* Dover, New York, 1995. Original edition: Springer, 1968.