Combinatorial Proof Semantics

DOMINIC J. D. HUGHES
Stanford University

The paper *Proofs Without Syntax* [*Annals of Mathematics*, to appear] introduced the notion of a *combinatorial proof* for classical propositional logic. The present paper uses combinatorial proofs to define a semantics for classical propositional sequent calculus, an inductive translation from sequent proofs to combinatorial proofs. The semantics is abstract and efficient: abstract in the sense that it identifies many sequent proofs, and efficient in the sense that combinatorial proofs are polynomial-time checkable and the inductive translation is polynomial.

1 Introduction

This paper aims to solve the following problem for classical propositional sequent calculus:

IDEAL SEMANTICS PROBLEM

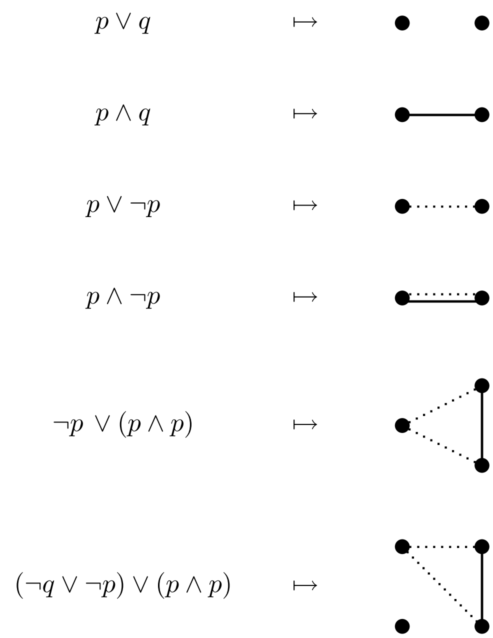
> *Find an efficient representation of classical propositional sequent calculus proofs which identifies as many proofs as possible.*

We shall formulate *efficiency* to preclude contrived representations which identify all proofs of a given sequent, such as the function mapping every proof of a sequent Γ to the truth table of Γ, or the constant function mapping every proof to the empty set. By "*identify many proofs*" we shall mean, for example, that the following two proofs, which differ only in the order of right-conjunction and left-weakening, have the same representation.

$$\dfrac{\dfrac{p \vdash p \qquad p \vdash p}{p \vdash p \wedge p}\,\wedge}{q, p \vdash p \wedge p}\,\mathsf{w} \qquad\qquad \dfrac{\dfrac{p \vdash p}{q, p \vdash p}\,\mathsf{w} \qquad \dfrac{p \vdash p}{q, p \vdash p}\,\mathsf{w}}{q, p \vdash p \wedge p}\,\wedge$$

We shall represent a sequent proof as a *combinatorial proof*, a notion introduced in [Hug04]. The presentation of combinatorial proofs in the current paper is more abstract.

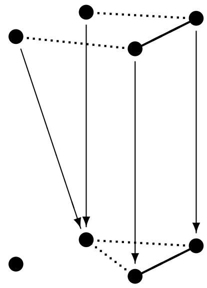
Combinatorial propositions. We shall represent a sequent abstractly as a ***bigraph***, by which we mean a simple undirected graph with two edge sets, rather than the usual one. The leaves of the parse tree of a sequent or formula become the vertices of its bigraph, the first edge set represents conjunctive relationships between leaves, and the second represents duality between leaves. Rather than getting bogged down in a formal definition here in the Introduction, we sketch the idea behind the translation with a progression of simple examples:



The bigraph of each formula is shown to its right. The first edge set, representing conjunctive relationships, is shown with solid edges; the second edge set, representing duality, is shown with dotted edges. The bigraph of first example $p \vee q$ has no edges: the p and q are neither conjunctively related nor dual. The bigraph of $p \wedge q$ has a solid edge, since the p and q are conjunctively related. The bigraph of $p \vee \neg p$ has a dotted edge, since p and $\neg p$ are dual. The bigraph of $p \wedge \neg p$ has both a solid edge and a dotted edge, since the p and the $\neg p$ are conjunctively related and dual. And so on.

We shall call any bigraph derivable from a sequent or formula a ***combinatorial proposition***. Combinatorial propositions are characterisable non-inductively as the P_4-free bigraphs whose second (dotted, duality) edge set has no odd cycle.

Bigraph homomorphisms. Our semantics represents a sequent proof as a bigraph homomorphism. For example, the two proofs in previous column both translate to the following bigraph homomorphism:



The vertex function, from a four-vertex bigraph to a four-vertex bigraph, is shown by the downward arrows. Note that this is indeed a bigraph homomorphism, since it preserves both $\wedge$-edges (solid) and $\neg$-edges (dotted).

The lower bigraph, the target, is the bigraph of the concluding sequent $q, p \vdash p \wedge p$. (The isolated vertex comes from q, the $\wedge$-edge (solid) models the conjunction $p \wedge p$, and the two $\neg$-edges (dotted) model the duality between the p on the left of the turnstile and the two p's on the right of the turnstile. Reading $q, p \vdash p \wedge p$ as the formula $(\neg q \vee \neg p) \vee (p \wedge p)$, this translation to a bigraph should be familiar as the last of the six examples on the previous page.) The four vertices of the upper bigraph, the source of the bigraph homomorphism, come from the four occurrences of p in the two axioms at the top of the proofs.

Combinatorial proofs. Write $G(\Gamma)$ for the bigraph of a sequent or formula Γ. We shall define a (cut-free) ***combinatorial proof*** of Γ as a bigraph homomorphism into $G(\Gamma)$ satisfying two conditions. One demands that the source of the homomorphism is *multiplicative*, a property related to multiplicative proof nets [Gir87, DR89, Ret03], and the other demands that the homomorphism is *structural*, a combinatorial analogue of a composite of structural sequent calculus rules. Any bigraph homomorphism derived from a sequent calculus proof satisfies the conditions, hence is a combinatorial proof. For example, the bigraph homomorphism displayed above, the translation of the two sequent proofs on page 1, is a (cut-free) combinatorial proof of $q, p \vdash p \wedge p$.

Efficiency. Returning to our formulation of the IDEAL SEMANTICS PROBLEM at the beginning of the paper, combinatorial proofs are an *efficient representation* of sequent proofs in the following sense.

EFFICIENT REPRESENTATION

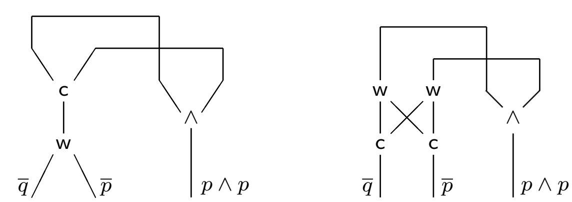
- *Correctness of combinatorial proofs is checkable in polynomial time.*
- *The function from sequent proofs to combinatorial proofs is polynomial time.*

A proof complexity theorist would say that combinatorial proofs constitute a formal *proof system* [CR79] which polynomially simulates propositional sequent calculus. (For a very readable introduction to propositional proof complexity, see [Urq95].) This definition of efficient representation precludes the contrived truth table function and constant function mentioned at the beginning of the Introduction. By incorporating complexity, the IDEAL SEMANTICS PROBLEM provides a very concrete formulation of the problem of finding a semantics for classical proofs.

Improvement on proof nets and linkings. Like combinatorial proofs, *proof nets* [Gir91] (clarified in [Rob03]) are also an efficient representation of classical propositional proofs. However, unlike combinatorial proofs, proof nets fail to identify the two sequent proofs

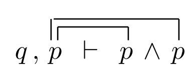
$$\dfrac{\dfrac{p \vdash p \qquad p \vdash p}{p \;\vdash p \wedge p}\,\wedge}{p, q \;\vdash p \wedge p}\,\mathsf{w} \qquad\qquad \dfrac{\dfrac{p \vdash p}{p, q \;\vdash p}\,\mathsf{w} \qquad \dfrac{p \vdash p}{p, q \;\vdash p}\,\mathsf{w}}{p, q \;\vdash p \wedge p}\,\wedge$$

considered earlier. Both proofs map to the combinatorial proof depicted at the top of the previous column. The respective (one-sided) proof nets are:



The failure to identify these proofs, and similar failures, forces the use of rewiring rules (coherence laws) on proof nets for cut elimination: see Table 8 of [FP06]. On combinatorial proofs, each of the rewiring rules in Table 8 becomes an equality.

One way to identify many sequent proofs is to follow the standard approach to multiplicative proof nets [Gir87]: simply trace the axioms down to links on the conclusion. The two sequent proofs above translate to the following linking:



This approach is studied in detail in [LS05]. Linkings fail to form a propositional proof system, since correctness is not polynomial.[1] Indeed, verifying a linking on Γ is no faster than verifying Γ itself, so linkings are redundant in the same way that an explicit truth table is redundant. Linkings fail to

[1]Were a polynomial-time correctness criterion ever to be found, it would yield a propositional proof system in which every tautology T can be verified by a certificate of size polynomial in the size in T, implying the remarkable complexity result $\mathsf{NP} = \mathsf{coNP}$ [CR79].

be an efficient representation of propositional sequent calculus proofs.

Combinatorial proofs are at an abstraction sweet spot between proof nets and linkings: they identify more proofs than proof nets, but not to the point of failing to be a proof system, as is the case for linkings.

Sequentialisation theorem. Analogous to Girard's sequentialisation theorem for multiplicative proof nets [Gir87], in Section 6 we obtain a sequentialisation theorem for a (complete) subclass of combinatorial proofs, the *binary* combinatorial proofs.

Cut elimination. Standard Gentzen-style cut elimination for sequent calculus adapts directly to proof nets (see [Rob03, FP06]). Just as Gentzen's original procedure is littered by transpositions of structural rules, the adapted procedure for proof nets is littered by rewiring rules [FP06, Table 8]. Combinatorial proofs obviate the rewiring rules, since the proof nets on either side of a rule are represented by the same combinatorial proof. It would be interesting to adapt Gentzen's cut elimination procedure to combinatorial proofs, in the hope of distilling the essence of the procedure, untainted by transpositions of structural rules or rewiring rules.

2 Preliminaries

Formulas. Fix a set $\mathcal{V}$ of ***variables***. A ***formula*** is any expression generated freely from variables by the binary operations ***and*** $\wedge$, ***or*** $\vee$, and ***implies*** $\Rightarrow$, and the unary operation ***not*** $\neg$. A ***valuation*** is a function $f : \mathcal{V} \to \{0,1\}$. Write $\hat{f}$ for the extension of a valuation f to formulas defined by $\hat{f}(\neg\phi) = 1-\hat{f}(\phi)$, $\hat{f}(\phi\wedge\rho) = \min\{\hat{f}(\phi),\hat{f}(\rho)\}$, $\hat{f}(\phi\vee\rho) = \max\{\hat{f}(\phi),\hat{f}(\rho)\}$, $\hat{f}(\phi\Rightarrow\rho) = \hat{f}((\neg\phi)\vee\rho)$. A formula ϕ is ***true***, or ***valid***, or a ***tautology***, if $\hat{f}(\phi)=1$ for all valuations f. Variables $p \in \mathcal{V}$ and their negations $\overline{p} = \neg p$ are ***literals***, and we say that p and $\overline{p}$ are ***dual***.[2]

Graphs. An ***edge*** on a set V is a two-element subset of V. A ***graph*** (V,E) is a finite set V of ***vertices*** and a set E of edges on V. Write $V(G)$ and $E(G)$ for the vertex set and edge set of a graph G, respectively, and vw for $\{v,w\}$. The ***complement*** of (V,E) is the graph (V,E^{c}) with $vw \in E^{\mathsf{c}}$ iff $vw \notin E$. The ***union*** $G \vee G'$ of graphs $G=(V,E)$ and $G'=(V',E')$ with no common vertex is $(V\cup V', E\cup E')$ and the ***join*** $G\wedge G'$ is $(V\cup V', E\cup E'\cup\{vv' : v\in V, v'\in V'\})$.

[2] To streamline our presentation we have excluded the constants 0 and 1 from the definition of formula. To recover constants, simply encode 0 as $p\wedge\overline{p}$ and 1 as $p\vee\overline{p}$, where p is a fresh variable for each occurrence of 0 and 1.

The ***empty*** graph is the graph with no vertices. A graph is ***disconnected*** if it is a union of non-empty graphs, otherwise it is ***connected***. A ***component*** is a maximal non-empty connected subgraph.

A graph (V,E) is a ***cograph***, or P_4***-free***, if V is non-empty and for any distinct $v,w,x,y \in V$, the restriction of E to edges on $\{v,w,x,y\}$ is not $\{vw,wx,xy\}$ (see *e.g.* [BLS99]). A graph G is ***bipartite*** if it has no odd cycle: whenever $v_1,\ldots,v_n$ are distinct vertices in G with n odd and $v_iv_{i+1} \in E(G)$ for $1 \le i < n$, then $v_nv_1 \notin E(G)$.

A ***homomorphism*** $h : G \to G'$ is a function $h : V(G) \to V(G')$ such that $vw \in E(G)$ implies $h(v)h(w) \in E(G')$. An ***isomorphism*** is a bijective homomorphism whose inverse function is also a homomorphism. Two graphs are ***isomorphic*** if there is an isomorphism between them.

A vertex set $C \subseteq V(G)$ is a ***clique*** if $vw \in E(G)$ for all distinct v and w in C.

3 Combinatorial propositions and truth

Bigraphs. A ***bigraph*** $G = (V,\wedge,\neg)$ is a finite set V of vertices together with sets $\wedge$ and $\neg$ of edges on V. The notation for the edge sets is chosen with a view to interpreting formulas as bigraphs. Write $G_\wedge$ and $G_\neg$ for the $\wedge$***-graph*** $(V,\wedge)$ and $\neg$***-graph*** $(V,\neg)$ of G, respectively. A $\wedge$***-edge*** of G is an edge of $G_\wedge$, and a $\neg$***-edge*** of G is an edge of $G_\neg$.

Given a formula ϕ, write $G(\phi)$ for the bigraph obtained from ϕ as follows. Without loss of generality, assume ϕ is generated by $\wedge$ and $\vee$ from literals (by de Morgan duality, $\theta \Rightarrow \psi \;\mapsto\; (\neg\theta\vee\psi)$ and $\neg\neg\theta \;\mapsto\; \theta$). The vertices of $G(\phi)$ are the leaves of the parse tree of ϕ, a pair of vertices vw is a $\wedge$-edge iff the smallest subformula of ϕ containing v and w is a conjunction, and vw is a $\neg$-edge iff v and w are labelled by dual literals. Six examples can be found on page 1.

A ***homomorphism*** $h : G \to G'$ between bigraphs is a function $h : V(G) \to V(G')$ which preserves $\wedge$-edges and $\neg$-edges, *i.e.*, which is simultaneously a graph homomorphism $G_\wedge \to G'_\wedge$ and $G_\neg \to G'_\neg$. An ***isomorphism*** is a bijective homomorphism whose inverse function is also a homomorphism. Two bigraphs are ***isomorphic*** if there is an isomorphism between them.

Given bigraphs $G = (V,\wedge,\neg)$ and $G' = (V',\wedge',\neg')$ with no common vertex, the ***union*** $G\vee G'$ is $(V\cup V', \wedge\cup\wedge', \neg\cup\neg')$ and the ***join*** $G\wedge G'$ is $G\vee G'$ together with a new $\wedge$-edge between every vertex of G and every vertex of G'. Join adds no $\neg$-edges.

Combinatorial propositions. A bigraph G is a ***combinatorial proposition*** if $G_\wedge$ and $G_\neg$ are P_4-free and $G_\neg$ is bipartite.

PROPOSITION 1 *A bigraph is a combinatorial proposition iff it is derivable from a formula.*

Proof. P_4-free graphs are precisely the graphs generated from individual vertices by union and join [BLS99, §11.3]. See Section 4 of [Hug04]. □

Combinatorial truth. Translating a formula to a combinatorial proposition forgets the names of literals. For example, each of $p \vee q$, $q \vee \neg p$, $\neg p \vee \neg q$ and $q \vee \neg r$ translates to the same combinatorial proposition $\bullet\ \ \bullet$. However, as we shall see in the definitions and lemma below, no information about validity is lost.

A ***clause*** in a graph is maximal set of vertices not containing an edge. A ***clause*** in a bigraph G is a maximal set of vertices not containing a $\wedge$-edge (*i.e.*, a clause in $G_\wedge$). A clause in G is ***true*** if it contains a $\neg$-edge, and G is true (or ***valid***) if each of its clauses is true.

PROPOSITION 2 *A formula is true iff its bigraph is true.*

Proof. Exhaustively apply distributivity $\theta \vee (\psi_1 \wedge \psi_2) \rightarrow (\theta \vee \psi_1) \wedge (\theta \vee \psi_2)$ to the formula ϕ, modulo associativity and commutativity of $\wedge$ and $\vee$, yielding a conjunction ϕ' of syntactic clauses (disjunctions of literals). The lemma is immediate for ϕ' since its bigraph $G(\phi')$ is a join of clauses together with additional $\neg$-edges, and $G\big(\theta \vee (\psi_1 \wedge \psi_2)\big)$ is true iff $G\big((\theta \vee \psi_1) \wedge (\theta \vee \psi_2)\big)$ is true since for non-empty graphs G_1 and G_2, a clause of $G_1 \vee G_2$ (resp. $G_1 \wedge G_2$) is a clause of G_1 and (resp. or) a clause of G_2. □

4 Structural homomorphisms

The key to our semantics of sequent proofs is to model structural rules very precisely as certain bigraph homomorphisms, called *structural homomorphisms*. In this section we prove the *Structural Characterisation Theorem*, which takes the following form:

> *One formula is derivable from another by structural rules iff there is a structural homomorphism between their bigraphs.*

In Section 5 we shall define a combinatorial proof as a structural homomorphism from a suitable source. The Structural Characterisation Theorem is the key to the Sequentialisation Theorem, which in turn yields the Soundness & Completeness Theorem.

Skeletons. A ***skeletal formula***, or ***skeleton***, is a formula generated by $\wedge$ and $\vee$ from the symbol $\bullet$, for example, $\bullet \wedge (\bullet \vee \bullet)$. Write $C(s)$ for the cograph (P_4-free graph) of a skeleton s, obtained by viewing $\bullet$ as a vertex and $\vee$ and $\wedge$ as union and join. Write $\simeq$ for skeleton equality modulo associativity and commutativity of $\wedge$ and $\vee$, and write $\cong$ for graph isomorphism. The following proposition is immediate.

PROPOSITION 3 $s \simeq t$ *iff* $C(s) \cong C(t)$, *for all skeletons* s *and* t.

One skeleton ***structurally implies*** another if the latter is derivable from the former by associativity and commutativity of $\wedge$ and $\vee$ together with the following rewrites applied to subformulas (subskeletons):

$$\begin{array}{rcll} s \vee s & \mapsto & s & \textit{(contraction)} \\ s & \mapsto & s \vee t & \textit{(weakening)} \end{array}$$

where t is an arbitrary skeleton. For example, the skeleton $(\bullet \vee \bullet) \vee (\bullet \wedge \bullet)$ structurally implies $\big((\bullet \vee \bullet) \wedge \bullet\big) \vee \bullet$:

$$\begin{array}{rcll} (\bullet \vee \bullet) \vee (\bullet \wedge \bullet) & \rightarrow & \bullet \vee (\bullet \wedge \bullet) & \text{(contract)} \\ & \rightarrow & (\bullet \wedge \bullet) \vee \bullet & \text{(commute)} \\ & \rightarrow & \big((\bullet \vee \bullet) \wedge \bullet\big) \vee \bullet & \text{(weaken)} \end{array}$$

Below we shall characterise structural implication semantically, proving that a skeleton s structurally implies a skeleton t iff there is a particular kind of graph homomorphism $C(s) \rightarrow C(t)$, called a *skew fibration*.

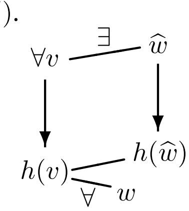
4.1 Cograph contraction and weakening

Given a cograph C, a ***skeleton of*** C is any skeleton s such that $C(s) \cong C$. Since cographs are generated from individual vertices by union and join [CLS81], every cograph has a skeleton, which by Proposition 3 is uniquely determined up to associativity and commutativity of $\wedge$ and $\vee$. This allows us to abandon skeletons in favour of working directly with cographs.

For any skeleton s and skeleton s' resulting from a weakening rewrite applied inside s, the canonical cograph homomorphism $C(s) \rightarrow C(s')$ is a ***weakening***. Similarly, for any skeleton s and skeleton s' resulting from a contraction rewrite applied inside s, the canonical cograph homomorphism $C(s) \rightarrow C(s')$ is a ***contraction***. For convenience, we shall also consider every graph isomorphism to be both a contraction and a weakening.

4.2 Skew fibrations

We recall the following definition from [Hug04, §3]. A graph homomorphism $h : G \rightarrow G'$ is a ***skew fibration*** if for all $v \in V(G)$ and $h(v)w \in E(G')$ there exists $v\widehat{w} \in E(G)$ with $h(\widehat{w})w \notin E(G')$.



The vertex $\widehat{w}$ is a ***skew lifting*** of w from v. If we demand $h(\widehat{w}) = w$ and uniqueness of $\widehat{w}$, then we have a standard graph fibration (simultaneously a special case of a topological fibration and a categorical fibration [Hug04, §3]).

Ignoring the dotted edges, the homomorphism displayed on page 2 is a skew fibration.

It is immediate from the statement of the definition that checking a function $h : V(G) \to V(G')$ is a skew fibration is $O(|G| \times \delta(G') \times \delta(G))$ where $|G|$ is the size of G (the number of vertices in G) and $\delta(H)$ denotes the maximum degree of a vertex in H. (The degree of a vertex is the number of edges containing it.) Thus the correctness of a skew fibration $G \to G'$ can be checked in polynomial time in the sizes of G and G'.

4.3 The Contraction-Weakening Theorem

A graph homomorphism $h : G \to G'$ ***preserves maximal cliques*** if for every maximal clique K of G, the image $h(K) = \{h(v) : v \in K\}$ is a maximal clique of G'.

THEOREM 1 (CONTRACTION-WEAKENING) *Let h be a homomorphism between cographs. The following are equivalent.*

(1) *h is a composite of contractions and weakenings.*

(2) *h preserves maximal cliques.*

(3) *h is a skew fibration.*

Proof. (1)⇒(2). It is easy to verify that any contraction or weakening preserves maximal cliques (by considering the underlying skeletal formulas). Maximal clique-preserving homomorphisms compose.

(2)⇒(3). A relatively routine graph-theoretic exercise.

(3)⇒(1). This is the tricky part of the theorem. The basis of our argument is lifted from the proof of the Combinatorial Soundness Theorem [Hug04, §5]. That proof iteratively decomposes a skew fibration using ***shallowness*** (the property that the inverse of every component is connected) and surjectivity, via Lemmas 5 and 6 of [Hug04]. The conversion to shallowness is readily observed to be a factorisation through post-composed contractions, and conversion to a surjection is a factorisation through post-composed weakenings (a result proved in Lemma 1 below). □

A skew fibration need not preserve maximal cliques if its target is not a cograph. Let C_n denote the n-vertex cycle, and let C_5^+ be C_5 with an additional edge. Inclusion $C_4 \to C_5^+$ is a skew fibration, but fails to preserve maximal cliques.

One way to interpret (2)⇔(3) in the Theorem is as follows: between cographs, checking max-clique preservation, which at first sight may seem exponential, can in fact be checked in polynomial time.

One proof obligation remains. The conversion of a skew fibration h to a surjection in the proof of the Combinatorial Soundness Theorem in [Hug04] post-composes with a full inclusion homomorphism i, where i is ***full*** if vw is an edge whenever $i(v)i(w)$ is an edge. This inclusion i inherits from h the property of being a skew fibration (a simple graph theoretic exercise). The proof of the Contraction-Weakening Theorem relied on the following.

LEMMA 1 *Any full inclusion homomorphism between cographs which is a skew fibration is a composite of weakenings.*

Proof. Let $i : C \to D$ be the inclusion. We proceed by induction on the size of D. If D is a vertex, the result is trivial. (Remember that every isomorphism is considered to be a weakening.)

Suppose $D = D_1 \wedge D_2$. Since i is a full inclusion, $C = C_1 \wedge C_2$ with cograph full inclusions $i_\alpha : C_\alpha \to D_\alpha$ restricted from i ($\alpha = 1, 2$), which are skew fibrations by Lemma 2 of [Hug04]. Since i is a skew fibration, each C_α is non-empty, hence is a cograph. By induction each i_α is a weakening composite, hence i is a weakening composite.

A similar argument applies to the case $D = D_1 \vee D_2$. □

4.4 The Structural Characterisation Theorem

To ***unify*** a combinatorial proposition P is to add $\neg$-edges to it so as to produce another combinatorial proposition P'. The canonical bigraph homomorphism $P \to P'$ is a ***unification***.

Let ϕ be a formula and ϕ' the result of applying a contraction rewrite to a subformula of ϕ, *i.e.*, replacing $\theta \vee \theta$ by θ somewhere in ϕ. The canonical bigraph homomorphism $C(\phi) \to C(\phi')$ is a ***contraction***. If ϕ' instead results by a weakening rewrite, *i.e.*, replacing θ by $\theta \vee \chi$ somewhere in ϕ for some formula χ, the canonical homomorphism is a ***weakening***.

For convenience, every isomorphism between combinatorial propositions will be considered simultaneously a unification, a contraction and a weakening.

A ***structural homomorphism*** $h : G \to G'$ is a bigraph homomorphism which is a skew fibration on the underlying $\wedge$-graphs, *i.e.*, the graph homomorphism $h_\wedge : G_\wedge \to G'_\wedge$ defined by $h_\wedge(v) = h(v)$ is a skew fibration. The bigraph homomorphism displayed on page 2 is a structural homomorphism.

THEOREM 2 (STRUCTURAL CHARACTERISATION) *A bigraph homomorphism between combinatorial propositions is a composite of unifications, contractions and weakenings iff it is a structural homorphism.*

Since the $\neg$-graph of a combinatorial proposition has a simple structure (bipartite cograph), the theorem derives without too much difficulty from the Contraction-Weakening Theorem.

5 Combinatorial proofs

A set W of vertices in a graph G ***induces a matching*** if it is non-empty and for all $w \in W$ there is a unique $w' \in W$ such that $ww' \in E(G)$. A set of vertices in a bigraph G ***induces a bimatching*** if it simultaneously induces a matching in $G_\wedge$ and in $G_\neg$. A combinatorial proposition is ***multiplicative*** if it has no induced bimatching and every vertex is in a $\neg$-edge. This property is related to multiplicative proof nets [Gir87, DR89, Ret03]. We shall see later that checking a combinatorial proposition is multiplicative can be done in polynomial time in its size.

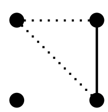
COMBINATORIAL PROOF (CUT-FREE DEF.)

*A cut-free **combinatorial proof** of a combinatorial proposition P is a structural homomorphism from a multiplicative combinatorial proposition to P.*

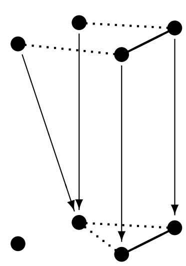
A cut-free combinatorial proof of a formula or sequent is a combinatorial proof of its combinatorial proposition.

The source of a combinatorial proof is its ***multiplicative core***, or simply ***core***.

For example, let P be the following combinatorial proposition (the last bigraph example on page 1):

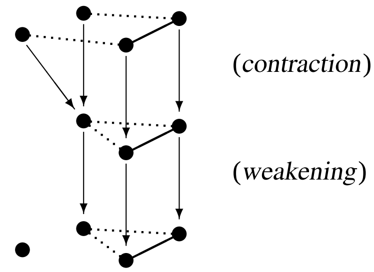


Here is a combinatorial proof of P (from page 2):



It is a combinatorial proof of the sequent $q, p \vdash p \wedge p$, and of the formula $(\overline{q} \vee \overline{p}) \vee (p \wedge p)$, since both translate to P. Note that the core is indeed multiplicative: all four vertices of the upper graph are in a $\neg$-edge (dotted), and no vertex set induces a bimatching. Also note that the arrows do indeed define a bigraph homomorphsim which is a structural homomorphism, *i.e.*, a skew fibration with respect to the solid $\wedge$-edges, or equivalently (by (2)$\Leftrightarrow$(3) in the Contraction-Weakening theorem), maximal $\wedge$-clique preserving.

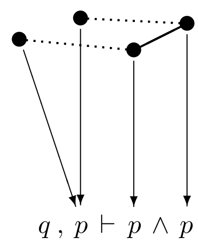
Here is an illustration of the Structural Characterisation Theorem, decomposing the structural homomorphism into a contraction followed by a weakening:

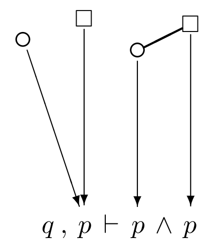


It also decomoses as a weakening followed by a contraction.

5.1 Colour depiction

Since the vertices of the bigraph $G(\phi)$ are the leaves of the parse tree of ϕ, we can abbreviate a combinatorial proof of ϕ by targetting the arrows into ϕ. For example, the above combinatorial proof of $q, p \vdash p \wedge p$ becomes as shown below-left:





Depicting each $\neg$-edge as a colour class (using colours $\circ$ and $\square$), we obtain the representation above-right. Finally, by drawing the coloured vertices directly over the leaves of the sequent, we can omit the arrows:

$$q\ ,\ \overset{\circ\square}{p} \vdash \overset{\circ}{p} \wedge \overset{\square}{p}$$

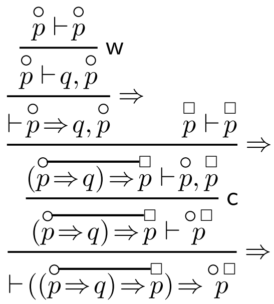
Later when we come to define the translation of a sequent proof into a combinatorial proof, the compact colour notation will be convenient for presenting an inductive translation. For example, recall the sequent proofs of $q\ ,\ p \vdash p \wedge p$ in the Introduction. Each becomes a tree of operations on combinatorial proofs, starting from the identity homomorphism on the bigraph $\bullet \cdots \bullet$ ($\overset{\circ}{p} \vdash \overset{\circ}{p}$ in a colour depiction) whose end product is the combinatorial proof above:

$$\dfrac{\dfrac{\overset{\circ}{p} \vdash \overset{\circ}{p} \qquad \overset{\square}{p} \vdash \overset{\square}{p}}{\overset{\circ\square}{p} \ \vdash \overset{\circ}{p} \wedge \overset{\square}{p}}\,\wedge}{q\ ,\ \overset{\circ\square}{p} \ \vdash \overset{\circ}{p} \wedge \overset{\square}{p}}\,\mathsf{w} \qquad\qquad \dfrac{\dfrac{\overset{\circ}{p} \vdash \overset{\circ}{p}}{q\ ,\ \overset{\circ}{p} \vdash \overset{\circ}{p}}\,\mathsf{w} \qquad \dfrac{\overset{\square}{p} \vdash \overset{\square}{p}}{q\ ,\ \overset{\square}{p} \vdash \overset{\square}{p}}\,\mathsf{w}}{q\ ,\ \overset{\circ\square}{p} \ \vdash \overset{\circ}{p} \wedge \overset{\square}{p}}\,\wedge$$

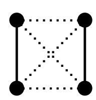
Below is another example in colour notation, a combinatorial proof of Peirce's law $((p \Rightarrow q) \Rightarrow p) \Rightarrow p$.

$$((\overset{\circ}{p} \Rightarrow q) \Rightarrow \overset{\square}{p}) \Rightarrow \overset{\circ\square}{p}$$

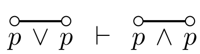
Here is the inductive translation of a sequent proof into this combinatorial proof:



In the examples above, each colour class was binary (*i.e.*, had two vertices), coming from a single edge of the ¬-graph which was isolated (in the sense that it did not share a vertex with any other ¬-edge). More generally, when the ¬-edges are not isolated, we depict each connected component (a complete bipartite subgraph) of the ¬-graph as a colour class. For example, the bigraph G of the sequent $p \vee p \vdash p \wedge p$ is shown below:

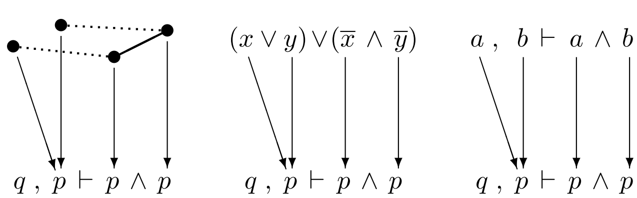


The identity bigraph homomorphism $i : G \to G$ is a combinatorial proof: G is multiplicative (since it has no induced bimatching, and every vertex is in some ¬-edge) and every identity is a graph fibration (hence a skew fibration). Depicting the combinatorial proof i in colour notation, we obtain an example with a four-vertex colour class as the multiplicative core, since the ¬-graph of G has a single (complete bipartite) component:



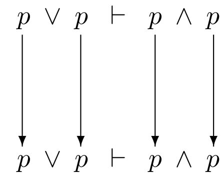
5.2 Semi-combinatorial depiction

Below-left we have reproduced one of the intermediate steps of the colour abbreviation in our first example.



Since every combinatorial proposition is the translation of a formula, we can choose a formula to represent the multiplicative core, as shown above-centre. Above-right, we implement the same idea with a sequent instead.

The identity combinatorial proof $G \to G$ for G the bigraph of $p \vee p \vdash p \wedge p$, drawn in colour mode in the previous subsection, can be presented in semi-combinatorial fashion as follows:



5.3 Syntactic presentation

Encoding the structural homomorphism as a sequence of integers (where n in i^{th} position denotes that the i^{th} leaf above maps to the n^{th} leaf below), we obtain a syntactic notation:

$$\frac{a\,,\,b \vdash a \wedge b}{q\,,\,p \vdash p \wedge p}\,2234$$

Alternatively,

$$\frac{a_2\,,\,b_2 \vdash a_3 \wedge b_4}{q\,,\,p \vdash p \wedge p}$$

or perhaps

$$a\,,\,b \vdash a \wedge b \quad \underset{2234}{\triangleright} \quad q\,,\,p \vdash p \wedge p$$

6 Sequentialisation Theorem

To ***contract*** (resp. ***weaken***, ***unify***) a combinatorial proof is to post-compose it with a contraction (resp. weakening, unification). (See Section 4.4.)

A ***portion*** of a graph G is a union of components in G (possibly empty). Given portions X of G and Y of H, the ***partial join*** of G and H at X and Y, denoted $G\,{}^{X}\!\wedge^{Y} H$, is obtained from the union $G \vee H$ by adding an edge between every vertex of X and every vertex of Y. Thus if either of X or Y is empty, $G\,{}^{X}\!\wedge^{Y} H$ is the union $G \vee H$, and if $X = G$ and $Y = H$ then $G\,{}^{X}\!\wedge^{Y} H$ is the join $G \wedge H$.

Given a graph homomorphism $g : G' \to G$ the ***lift*** of a portion X of G (with respect to g) is the portion of G' which maps into X (*i.e.*, the union of the components C of G' such that $h(V(C)) \subseteq V(X)$). Given graph homomorphisms $g : G' \to G$ and $h : H' \to H$ and portions X of G and Y of H, the ***partial join*** of g and h at X and Y, denoted $g\,{}^{X}\!\wedge^{Y} h$ is the homomorphism $g \cup h \ : \ G'\,{}^{X'}\!\wedge^{Y'} H' \to G\,{}^{X}\!\wedge^{Y} H$ where X' is the lift of X' and Y' is the lift of Y. The partial join is a ***fusion*** if X and Y are ***coherent***: their lifts X' and Y' are either both empty, or both non-empty.

A ***portion*** of a bigraph is a portion of its ∧-graph. Define the partial join of bigraphs by forming the partial join on the underlying ∧-graphs, carrying the ¬-edges along. Similarly, define the partial join of bigraph homomorphisms by forming the partial join with the underlying homomorphism of ∧-graphs. Define fusion as before (a partial join between coherent portions).

A combinatorial proposition is ***binary*** if every vertex is in exactly one $\neg$-edge, and a combinatorial proof is binary if its multiplicative core is binary. Write $\mathbb{T}$, called ***true***, for the identity homomorphism on the bigraph $\bullet\cdots\bullet$ (the combinatorial proposition with two vertices, no $\wedge$-edge and one $\neg$-edge).

THEOREM 3 (SEQUENTIALISATION) *A bigraph homomorphism is a cut-free binary combinatorial proof iff it is derivable from $\mathbb{T}$ by fusion, contraction, weakening and unification.*

Proof. Every derivable homomorphism is a combinatorial proof since $\mathbb{T}$ is a combinatorial proof and a fusion of combinatorial proofs is a combinatorial proof.

Conversely, by the Structural Characterisation Theorem (page 5), every combinatorial proof is derivable from (an identity homomorphism on) a binary combinatorial axiom by contraction, weakening and unification. Thus sequentialisation reduces to showing that every binary combinatorial axiom is derivable from $\mathbb{T}$ = $\bullet\cdots\bullet$ by partial join. This is by Lemma 8 of [Hug04], since a binary combinatorial axiom is a nicely coloured coloured cograph upon viewing $\neg$-edges as two-vertex colour classes, and partial join of bigraphs corresponds to fusion of coloured graphs in [Hug04]. □

Since fusion, contraction, weakening and unification all preserve validity of the conclusion (target) of a combinatorial proof, we have:

THEOREM 4 (CUT-FREE BINARY SOUNDNESS) *If a combinatorial proposition has a cut-free binary combinatorial proof, it is valid.*

We extend soundness to all combinatorial proofs in Section 8.

7 Semantics of sequent proofs

Without loss of generality (by preliminary translation if necessary), we work with classical propositional sequent calculus formulated as multiplicative linear logic [Gir87] together with contraction and weakening. Sequents (nonempty sequences $\phi_1, \ldots, \phi_n$ of formulas generated from literals by $\wedge$ and $\vee$) are proved using the following rules:[3]

$$\overline{p},\, p \qquad\quad \dfrac{\Gamma, \phi \qquad \psi, \Delta}{\Gamma, \phi \wedge \psi, \Delta}\,\wedge \qquad\quad \dfrac{\Gamma,\, \phi,\, \psi}{\Gamma,\, \phi \vee \psi}\,\vee$$

$$\dfrac{\Gamma,\, \phi,\, \psi,\, \Delta}{\Gamma,\, \psi,\, \phi,\, \Delta}\,\mathsf{x} \qquad\quad \dfrac{\Gamma,\, \phi,\, \phi}{\Gamma,\, \phi}\,\mathsf{c} \qquad\quad \dfrac{\Gamma,\, \phi}{\Gamma,\, \phi, \psi}\,\mathsf{w}$$

Here p ranges over variables, ϕ and ψ range over formulas, and Γ and Δ range over sequence of formulas.

We interpret the axiom $\overline{p}, p$ (for any variable p) as the combinatorial proof $\mathbb{T}$ (the identity homomorphism on the bigraph $\bullet\cdots\bullet$). Each single-hypothesis rule post-composes with a structural homomorphism in the obvious way: the $\vee$ and x rules by an isomorphism, the c rule by a contraction, and the w rule by a weakening and a possible unification (if ψ shares any variables with Γ or ϕ).

7.1 Garbage collection

The naive interpretation of the $\wedge$ rule is as a partial join operation on the bigraph homomorphism derived from the left branch and the bigraph homomorphism derived from the right branch, with portions corresponding to ϕ and ψ. However, this fails to preserve the skew fibration condition when a weak formula meets a strong formula in the rule (*i.e.*, when no variable traces down from an axiom in the left branch down to ϕ, but one or more variables can be traced onto ψ in the right branch, or vice versa). For example:

$$\dfrac{\dfrac{\overline{p},\, p}{\overline{p},\, p,\, q}\,\mathsf{w} \qquad\quad \overline{r},\, r}{\overline{p},\, p,\, q \wedge \overline{r},\, r}\,\wedge$$

The conjunction is between the weak q and the strong $\overline{r}$. With the naive interpretation of the $\wedge$ rule, the proof translates to the following tree of operations on bigraph homomorphisms (using the colour notation of Section 5.1):

$$\dfrac{\dfrac{\overset{\circ}{\overline{p}},\, \overset{\circ}{p}}{\overset{\circ}{\overline{p}},\, \overset{\circ}{p},\, q}\,\mathsf{w} \qquad\quad \overset{\square}{\overline{r}},\, \overset{\square}{r}}{\overset{\circ}{\overline{p}},\, \overset{\circ}{p},\, q \wedge \overset{\square}{\overline{r}},\, \overset{\square}{r}}\,\wedge$$

The final bigraph homomorphism is not a combinatorial proof, since it fails the skew fibration condition (since the leaf q does not have a skew lifting from the square over $\overline{r}$).

To address this situation we shall do some garbage collection when interpreting the $\wedge$ rule. We consider two approaches: (a) collect garbage on the sequent proof, before translation, to ensure that the above situation does not arise, and (b) perform the translation first, yielding a bigraph homomorphism h possibly failing to be a skew fibration, then garbage collect on h, deleting vertices in its source until reducing it to a skew fibration. The latter approach makes more natural identifications on sequent calculus proofs, and therefore we shall take it to be the definitive combinatorial proof semantics. The differences between the translations are interesting and subtle, and are the topic of a forthcoming companion paper.

Garbage collection before translation. With reference to the example above, if the portions corresponding to ϕ and ψ are coherent (*i.e.*, their lifts X and Y are both empty or

[3] As remarked in footnote 2, constants 0 and 1 can be encoded if desired. We add the cut rule in the next section.

both non-empty), we shall perform a partial join (which will be a fusion, because of the coherence); however, if the left lift X is empty while the right lift Y is not, we shall ignore the combinatorial proof from the right branch, and interpret the rule by weakening the combinatorial proof from the left branch (and similarly with the roles of X and Y reversed). Thus the sequent proof above translates to the following tree of operations, concluding with a well-defined combinatorial proof, satisfying the skew fibration condition:

$$\dfrac{\dfrac{\overset{\circ}{\overline{p}},\ \overset{\circ}{p}}{\overset{\circ}{\overline{p}},\ \overset{\circ}{p},\ q}\,\mathsf{W} \qquad \overset{\square}{\overline{r}},\ \overset{\square}{r}}{\overset{\circ}{\overline{p}},\ \overset{\circ}{p},\ q\wedge\overline{r},\ r}\,\wedge$$

Garbage collection after translation. Given a sequent calculus proof π of Γ, let $h_\pi : G' \to G$ be the bigraph homomorphism obtained by direct translation, ***i.e.***, with the naive interpretation of the $\wedge$ rule which never deletes one of its arguments. Thus the $2n$ vertices of G' come from the n axioms at the top of π. By the discussion at the beginning of this subsection, h in general fails to be a skew fibration. Write $|h_\pi|$ for the combinatorial proof obtained by exhaustively performing the following ***garbage collection*** on G': delete a vertex v if either (a) h lacks a skew lifting from v or (b) v is not in a $\neg$-edge. Garbage collection is easily seen to be confluent and terminating, and (by a simple induction) since h_π is translated from a sequent calculus proof, $|h_\pi|$ is non-empty.

Postponing garbage collection until after translation leads to more natural identifications on sequent calculus proofs, therefore we take this to be the definitive combinatorial proof semantics. The subtle difference between collecting garbage before and after translation is the subject of a sibling paper. In the simple example on $\overline{p}, p, q\wedge\overline{r}, r$ above, the final combinatorial proof is the same, irrespective of when garbage collection is performed.

7.2 Corollary: Completeness

The completeness of propositional sequent calculus combined with either of the above translations from a cut-free syntactic proof to a cut-free binary combinatorial proof yields:

THEOREM 5 (COMPLETENESS) *Every valid combinatorial proposition has a cut-free binary combinatorial proof.*

8 Combinatorial proofs with cuts

A ***cut*** is a formula of form $\psi \wedge \neg\psi$ (where we treat $\neg$ as a derived operation on a $\wedge\vee$-normal formula ψ). A ***cut-extension*** of a sequent Γ is any sequent $\Gamma, \chi_1, \ldots, \chi_n$ for $n \geq 0$ and cuts χ_i; correspondingly, the bigraph translation of $\Gamma, \chi_1, \ldots, \chi_n$ is a ***cut-extension*** of the bigraph translation $G(\Gamma)$ of Γ. A ***combinatorial proof*** of a sequent Γ (resp. combinatorial proposition G) is a cut-free combinatorial proof of any cut-extension of Γ (resp. G). Adding the following form of the cut rule

$$\frac{\Gamma, \psi \wedge \neg\psi}{\Gamma}$$

to our sequent calculus allows us to place all cut rules, without loss of generalty, at the end of the proof, hence the translation of a sequent calculus proof of Γ with n instances of the cut rule to a combinatorial proof of Γ with n cuts is immediate.

Since a cut extension of Γ is a tautology iff Γ is a tautology, we have the following corollary of cut-free binary soundness (Theorem 4):

COROLLARY 1 (BINARY SOUNDNESS) *If a combinatorial proposition has a binary combinatorial proof, it is valid.*

Finally, we obtain soundness for the full system of non-binary combinatorial proofs by translation into binary combinatorial proofs.

THEOREM 6 (SOUNDNESS) *If a combinatorial proposition has a combinatorial proof, it is valid.*

Proof. We map every non-binary combinatorial proof $h : G \to G'$ to a binary combinatorial proof with two-vertex cuts. Suppose $G_\neg$ has a connected component C with three or more vertices, necessarily a complete bipartite subgraph between m vertices in C_1 and n vertices in C_2. Let C' be a copy of C. Add C' (disjointly) to $G_\wedge$, interpreting the edges of C' (which had been $\neg$-edges) now as $\wedge$-edges; add m parallel (***i.e.***, disjoint) $\neg$-edges between C_1 in G and its copy C'_1 and n parallel $\neg$-edges between C_2 in G and its copy C'_2; add a two-vertex cut to G', with vertices $\{v_1, v_2\}$, and unify as necessary to preserve the property of being a combinatorial proposition (adding any requisite $\neg$-edges between the v_i and the vertices of G'); extend h to C' with $h(w) = v_i$ for all $w \in C'_i$. Iterate this process to obtain a binary combinatorial proof of G', with two-vertex cuts. $\square$

8.1 Modelling resolution proofs as cut-free combinatorial proofs.

The inverse of the translation in the proof above provides a model of propositional resolution proofs as cut-free combinatorial proofs, as follows. Given a resolution proof π of Γ, first translate it to a sequent calculus proof $S(\pi)$ of Γ with atomic cuts, then (using the translation defined in Section 7) translate $S(\pi)$ to a binary combinatorial proof $h : G \to G(\Gamma \vee \chi_1 \ldots \chi_n)$, where the χ_i are the atomic cut formulas obtained from the cut rules in $S(\pi)$. By the nature of the syntactic translation $S(-)$, the combinatorial proof h

will have the following property: at least one vertex of every $\neg$-edge in G maps to a leaf of Γ (rather than to a leaf of one the cuts χ_i). Thus we can delete each two-vertex cut C, and the vertices mapping to it under h, and replace it with $m \times n$ edges forming an bipartite complete graph in the $\neg$-graph of G, where m and n are the numbers of vertices of G mapped by h to the two vertices of C, respectively.

9 Complexity

Combinatorial proofs form a *proof system* [CR79], since correctness is polynomial time. We have already seen that checking a skew fibration is polynomial. Checking a binary combinatorial axiom is polynomial by standard breadth-first search on its modular decomposition tree [BLS99], and the conversion of a non-binary combinatorial axiom to a binary combinatorial axiom described just before the Soundness Theorem is polynomial-time.

The translation from sequent proofs is clearly polynomial-time if we represent the combinatorial proofs in syntactic form during the translation (as defined in Section 5.3). Via the translation, combinatorial proofs inherit the sequent calculus property of having no obvious superpolynomial lower bounds [Urq95]. Thus combinatorial proofs present an approach towards $\mathsf{NP} \overset{?}{=} \mathsf{coNP}$ and $\mathsf{P} \overset{?}{=} \mathsf{NP}$ which is logical, yet disencumbered of syntax.

References

[BLS99] Brandstädt, A., V. B. Le & J. P. Spinrad. *Graph Classes: A Survey. SIAM monographs on Discr. Math. & Applic.* 1999.

[CR79] Cook, S. A. & R. A. Reckhow. *The relative efficiency of propositional proof systems. J. Symb. Logic* 44(1) 1979 36–50.

[CLS81] Corneil, D. G., H. Lerchs & L. K. Stewart-Burlingham. *Complement reducible graphs. Discr. Appl. Math.* 3 1981 163–174.

[DR89] Danos, V. & L. Regnier. *The structure of multiplicatives. Archive for Math. Logic* 28 1989 181–203.

[FP06] Führmann, C. & D. Pym *Order-enriched categorical models of the classical sequent calculus Journal of Pure and Applied Algebra* 204(1) 21-78 2006.

[Gen35] Gentzen, G. *Untersuchungen über das logische Schließen I, II. Mathematische Zeitschrift* 39 1935 176–210 405–431. Translation in M.E. Szabo, *The Collected Papers of Gerhard Gentzen,* North-Holland 1969.

[Gir87] Girard, J.-Y. *Linear logic*. *Th. Comp. Sci.* 50 1987 1–102.

[Gir91] Girard, J.-Y. *A new constructive logic: classical logic*. *Math. Struc. Comp. Sci.* 1 1991 255-296.

[Hug04] Hughes, D. J. D. *Proofs Without Syntax. Annals of Mathematics,* to appear. Draft *http://boole.stanford.edu/˜dominic/papers/pws* . Submitted version of August 2004: *http://arxiv.org/abs/math/0408282* .

[LS05] Lamarche, F. & L. Straßburger *Naming Proofs in Classical Propositional Logic*. Proc. TLCA'05, *Lecture Notes in Computer Science* 3461 246–261 2005.

[Ret03] Retoré, C. *Handsome proof-nets: perfect matchings and cographs. Th. Comp. Sci.* 294 2003 473–488.

[Rob03] Robinson, E. P. *Proof Nets for Classical Logic. J. Logic & Computation* 13(5) 777–797 2003.

[Urq95] Urquhart, A. *The complexity of propositional proofs. Bull. Symb. Logic* 1(4) 425-467 1995.