Herbrand proofs and expansion proofs as decomposed proofs

BENJAMIN RALPH, *Department of Computer Science, University of Bath, Bath, Somerset, UK.*
E-mail: bdr25@bath.ac.uk

Abstract

The reduction of undecidable first-order logic to decidable propositional logic via Herbrand's theorem has long been of interest to theoretical computer science, with the notion of a Herbrand proof motivating the definition of expansion proofs. In this paper we construct simple deep inference systems for first-order logic, both with and without cut, such that 'decomposed' proofs—proofs where the contractive and non-contractive behaviour of the proof is separated—in each system correspond to either expansion proofs or Herbrand proofs. Translations between proofs in this system, expansion proofs and Herbrand proofs are given, retaining much of the structure in each direction.

1 Introduction

1.1 Herbrand's theorem

Much of the development of first-order proof theory was driven by Hilbert's Program, an attempt by the (largely overlapping) mathematical and philosophical community to rebuild faith in set theory, to retake 'the paradise that Cantor created for us' by formalizing and proving the consistency of infinitary mathematics by finitary means [22]. This is the context in which Herbrand's work was carried out—in essence, Herbrand's project was to find a concise representation of the content of first-order proofs that was truly first order, as opposed to merely propositional. Using this representation, he aimed to prove the consistency and completeness of first-order logic and can be seen as a close relative of Gödel's completeness theorem. The importance of such a project, especially to the then still embryonic field of theoretical computer science, is that while propositional logic is decidable, full first-order logic is not. With this in mind, Herbrand's theorem [20] can be seen as teasing out the kernel of undecidability from first-order logic.

The basic idea of Herbrand's theorem is perhaps best introduced by a simple example, one common in the literature. Take the first-order translation of the sentence 'There exists two irrational numbers a and b such that a^b is rational.', which we will abbreviate as $(\overline{\mathbb{Q}}(a) \wedge \overline{\mathbb{Q}}(b) \wedge \mathbb{Q}(a^b))$ (Figure 1). For an intuitionist, a proof of the statement would have to be a pair of rationals (a, b) that fit the bill. However, classical proofs can be more liberal: if for some finite list of pairs $(a_1, b_1), \ldots, (a_n, b_n)$ we can prove that $\bigvee_1^n(\overline{\mathbb{Q}}(a_i) \wedge \overline{\mathbb{Q}}(b_i) \wedge \mathbb{Q}(a_i^{b_i}))$ is a true sentence, then we have proved the original statement. It turns out that the second approach gives us a simpler proof than the first. For if we choose $a_1 = \sqrt{2}$, $b_1 = \sqrt{2}$, $a_2 = \sqrt{2}^{\sqrt{2}}$ and $b_2 = \sqrt{2}$, proving the required statement is simply shown to reduce to proving that $\sqrt{2}^{\sqrt{2}}$ is either rational or irrational.[1] This strategy of proving existential formula by expanding them into a finite disjunction of instantiations is the crux of Herbrand's theorem.

[1] In fact, Kuzmin proved $\sqrt{2}^{\sqrt{2}}$ to be transcendental [27]

Vol. 30, No. 8,
Advance Access Publication on 24 October 2020 doi:10.1093/logcom/exaa052

$$\begin{array}{ccc}
 & \exists a,b\in\mathbb{R}(\overline{\mathbb{Q}}(a)\wedge\overline{\mathbb{Q}}(b)\wedge\mathbb{Q}(a^b)) & \\
\swarrow & & \searrow \\
\exists a,b(\overline{\mathbb{Q}}(a)\wedge\overline{\mathbb{Q}}(b)\wedge\mathbb{Q}(a^b)) & \vee & \exists a,b(\overline{\mathbb{Q}}(a)\wedge\overline{\mathbb{Q}}(b)\wedge\mathbb{Q}(a^b)) \\
\downarrow & & \downarrow \\
\overline{\mathbb{Q}}(\sqrt{2})\wedge\overline{\mathbb{Q}}(\sqrt{2})\wedge\mathbb{Q}(\sqrt{2}^{\sqrt{2}}) & \vee & \overline{\mathbb{Q}}(\sqrt{2}^{\sqrt{2}})\wedge\overline{\mathbb{Q}}(\sqrt{2})\wedge\mathbb{Q}(2) \\
\downarrow & & \downarrow \\
\mathbb{Q}(\sqrt{2}^{\sqrt{2}}) & \vee & \overline{\mathbb{Q}}(\sqrt{2}^{\sqrt{2}})
\end{array}$$

FIGURE 1. A demonstration of $\exists a,b\in\mathbb{R}(\overline{\mathbb{Q}}(a)\wedge\overline{\mathbb{Q}}(b)\wedge\mathbb{Q}(a^b))$ in 'Herbrand' style.

Unfortunately, Herbrand's own statement of his theorem, let alone the proof, is notoriously 'hard to follow' [23] and many of his lemmas are incorrect [15]. Therefore, most contemporary treatments reformulate the material, both stating and proving the theorem using terminology and techniques not available to Herbrand.[2] For example, we have the following statement of Herbrand's theorem in a prominent, more recent exposition by Buss [12]:

THEOREM 1 (Herbrand's theorem).
A first-order formula A is valid if and only if A has a Herbrand proof. A Herbrand proof of A consists of a blueprenexification $A^\star$ of a redstrong $\vee$-expansion of A plus a magentawitnessing greensubstitution σ for $A^\star$.

Herbrand's theorem has also been stated and proven in a *deep inference* system, where we are able to freely compose derivations using propositional connectives and quantifiers [11, 16, 17]. The statement is as follows [10]:

THEOREM 2 (Herbrand's theorem).
For each proof of a formula S in system SKSgr there is a substitution σ, a propositional formula P, a context $Q\{\ \}$ consisting only of quantifiers and a *Herbrand proof*:

$$\begin{array}{l}
\quad \Pi\ \mathsf{KS}\cup\{\mathsf{ai}\uparrow\} \\
\forall\vec{x}P\sigma \\
\quad \Vert\ \{\mathsf{n}\downarrow\} \\
Q\{P\} \\
\quad \Vert\ \{\mathsf{gr}\downarrow\} \\
S' \\
\quad \Vert\ \{\mathsf{qc}\downarrow\} \\
S
\end{array}$$

[2] The author, with Jack Webb, is currently working on an article presenting a correct proof of the theorem that stays as close as possible to Herbrand's strategy, expanding on the co-author's masters thesis [37].

From these we can abstract a pattern of four key steps necessary for a Herbrand proof.

1. Expansion of existential subformulae.
2. Prenexification/elimination of universal quantifiers.
3. Term assignment.
4. Propositional tautology check.

This strategy is common to the two approaches. But we can also note the difference between the two formulations. One key difference between the two is that, while Buss' definition of a Herbrand proof is that it is a *sui generis* form of proof, not a particular class of proof in a particular proof formalism, whereas Brünnler's is merely a subclass of proofs in a particular deep inference system. In deep inference, each of the four conditions for the Herbrand proof correspond to certain first-order inference rules, rather than an *ad hoc* operation on a first-order formula. Is this not possible in the sequent calculus?

1.2 Herbrand's theorem as a decomposition theorem

The key to the difference between Buss's and Brünnler's Herbrand proofs can be found in one of the earliest deep inference papers, setting out two properties related to contraction that one might want for a classical proof system. The second property, the property of having a *decomposition theorem* [3, 5] is the following:

> *'Proofs can be separated into two phases (seen bottom-up): The lower phase only contains instances of contraction. The upper phase contains instances of the other rules, but no contraction. No formulae are duplicated in the upper phase.'* [9]

Brünnler shows that a standard sequent calculus proof system with multiplicative rules do not have a valid decomposition theorem. The suggested way round this restriction is to use systems with *deep contraction*. In fact, this restriction on sequent calculus systems is shown by McKinley in [28] to create a gap in Buss's proof of Herbrand's theorem in [12]. The faulty proof assumes that if one restricts contraction to only existential formulae, one retains completeness (assuming a multiplicative $\wedge R$ rule). That this is false can be seen by considering the sequent below, where the application of any multiplicative $\wedge R$ rule leads to an invalid sequent:

$$\vdash \forall xA \wedge \forall xB, \left(\exists x\bar{A} \vee \exists xB\right) \wedge \left(\exists x\bar{A} \vee \exists xB\right).$$

It is the inability of sequent systems to satisfy this property that ensures that Herbrand proofs can never be expressed as a subclass of sequent proofs. Moreover, the first stage of a Herbrand proof is duplicating existential formulae, which when translated into a bottom-up proof system is performed by contraction. Therefore, Herbrand proofs, in common with decomposed proofs, have contractions at the bottom of their proofs. Therefore, we can see Herbrand's theorem as the first-order instantiation of the more general proof theoretic procedure of decomposition.

1.3 Expansion proofs as decomposed proofs

Herbrand proofs are not the only way that Herbrand's theorem has been reinterpreted. Another strand of research was initiated by the definition of 'expansion proofs', a generalization of Herbrand's theorem to higher-order logics [30]. The idea is to enrich formulae, explicitly adding in substitution information as syntax, so that they contain the 'Herbrand' information' of a first-order proof of that formula. One intuitive way to think about expansion proofs is as Coquand-style games: ∃loise, who can choose terms at existential node, plays ∀belard, who chooses variables at universal nodes. Once

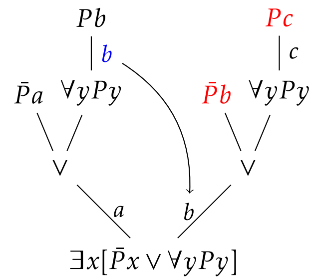


FIGURE 2. An expansion proof of the drinker's formula.

all the quantifiers are expanded, ∃loise wins if the resulting propositional formula is a tautology, ∀ otherwise. The first-order formula is true iff ∃loise has a winning strategy.

However, the game described above could only represent intuitionistically valid proofs. For classical proofs, we must give ∃loise the ability to 'backtrack', returning to any previously expanded existential node at any point to choose another term. The winning condition is now a disjunction over all of ∃loise's choices. Since ∃loise can only include free variables in her terms once ∀belard has played them, this gives her access to more winning strategies, matching the fact that more first-order sentences are true classically than intuitionistically.

As an example, we consider the drinker's formula, $\exists x \forall y[\bar{P}x \vee Py]$, as popularized by Smullyan: 'There is someone in the pub such that, if they are drinking, then everyone in the pub is drinking.' As the outermost quantifier is an existential, ∃loise moves first. At this point, there is no other move but to choose a closed term at random. ∀belard then chooses a variable to play—clearly he should not pick the same term as ∃loise. Assuming not, the resulting tautology would at this point be something along the lines of $[Pa \vee \bar{P}b]$, it seems as if ∃loise does not have a winning strategy. However, ∃loise is allowed to backtrack, choosing the variable ∀belard picks for her second existential witness. This time, whatever ∀belard picks, the disjunction over the two choices will be a tautology, say $(Pa \vee \bar{P}b) \vee (Pb \vee \bar{P}c)$.

In the original presentation of expansion proofs, Miller provides translations back and forth between his new formalism and the sequent calculus. However expansion proofs did not enjoy all the usual features of a proof system. Firstly, there is no account of the propositional aspect, just a tautology check. Obviously one could be given, but there is no natural analogue of expansion proofs for classical propositional logic. This isn't really a problem—the motivation behind expansion proofs is certification of first-order proofs, and using a first-order proof as a certificate for itself isn't of much use. Secondly, there is no means to compose proofs by cut and certainly no cut elimination. In fact, proving cut elimination for expansion proofs, or similar structures has been a relatively active topic of research in recent years. In [19], a system of 'proof forests' is presented, a graphical formalism of expansion proofs with cut and a local rewrite relation that performs cut elimination. Similar work has been carried out in [29]. A more categorical approach has also been given in [2].

Instead of a *sui generis* formalism for expansion proofs with cut, we show a class of deep-inference proofs that closely correspond to expansion proofs and another to expansion proofs with cuts, giving translations that are fully canonical in one direction, and partially in the other. Thus, expansion proofs represent certain key information contained in a proof, and can be used to guide normalization.

Thus, we have two different approaches to Herbrand's theorem: Herbrand proofs and expansion proofs. The first approach is from a more Hilbertian line of proof theory, with links often made to model theory and Gentzen's cut elimination results; the other integrated with newer traditions, such as game semantics and proof nets. However, apart from the definition of new inference rules, no real syntactic innovations are needed to situate both these approaches within open deduction proof systems, showcasing their capacity to internally describe a wide range of proof theoretic approaches.

1.4 Summary and open questions

In this paper, we present a study of Herbrand's theorem from the point of view of *open deduction*, a deep inference formalism. We describe a class of first-order proof, Herbrand proofs, and state and prove Herbrand's theorem for cut-free first-order proofs in open deduction as a transformation into a Herbrand proof. We introduce expansion proofs, and a class of first-order proof that corresponds closely to them: proofs in Herbrand Normal Form. We then show the translations between Expansion Proofs and proofs in Herbrand Normal Form, and cut elimination for expansion proofs with cuts, also known as proof forests, can used for an indirect deep inference cut elimination theorem. Most of the content of this paper is contained within my PhD Thesis [32]. Some proofs which are not so pertinent to the major themes of this paper have been omitted, all of these can be found in full in [32]. Finally, I would like to explicitly draw attention to a number of open questions and problems that are implicitly posed to the reader below:

- To what extent can we give native cut elimination proofs in SKSq that model the cut elimination relations for expansion proofs given in [19], [29] and [6]?
- As noted in [1], the rules of passage can be used to produce proofs of first-order sentences that are non-elementarily shorter than can be provided in a standard sequent system such as **LK**. How does this complexity analysis translate into the deep inference setting, where certain rules of passage are standard inference rules of even cut-free systems such as KSq?
- In 1900, Hilbert decided not to present what is now called his 'twenty-fourth problem' [24, 35, 36]. Put briefly, he was to ask the assembled mathematicians to '[f]ind criteria of simplicity or rather prove the greatest simplicity of given proofs.' To what extent do any of the first-order proof systems discussed in this paper satisfy Hilbert's desire for such criteria, and what evidence can be offered in support of such a claim?

2 Preliminaries on Open Deduction

As discussed above, we will work in the open deduction formalism. Open deduction differs from the sequent calculus in that we build up complex derivations with connectives and quantifiers in the same way that we build up formulae [17]. We can compose two derivations horizontally with $\vee$ or $\wedge$, quantify over derivations, and compose derivations vertically with an inference rule.

DEFINITION 3
An open deduction derivation is inductively defined in the following way:

- Every atom $Pt_1 \ldots t_n$ is a derivation, where P is an n-ary predicate, and t_i are terms. The units t and f are also derivations.

If $\begin{smallmatrix} A \\ \phi \| \\ B \end{smallmatrix}$ and $\begin{smallmatrix} C \\ \psi \| \\ D \end{smallmatrix}$ are derivations, then:

- • $\begin{array}{c} A \star C \\ \phi\star\psi \Vert \\ B \star D \end{array} = \begin{array}{c} A \\ \phi \Vert \\ B \end{array} \star \begin{array}{c} C \\ \psi \Vert \\ D \end{array}$ and $\begin{array}{c} QxA \\ Qx\phi \Vert \\ QxB \end{array} = Qx \left(\begin{array}{c} A \\ \phi \Vert \\ B \end{array} \right)$ are derivations.
- $\begin{array}{c} A \\ \chi \Vert \\ D \end{array} = \begin{array}{c} A \\ \phi \Vert \\ \rho \dfrac{B}{C} \\ \psi \Vert \\ D \end{array}$ is a derivation, if $\rho \dfrac{B}{C}$ is an instance of ρ.

When we write $\begin{array}{c} A \\ \phi \Vert S \\ B \end{array}$, it means that every inference rule in ϕ is an element of the finite set of inference rules S (we call S a *proof system*) or an equality rule.

REMARK 4
Formulae are just derivations built up with no vertical composition. Open deduction and the *calculus of structures* (the better known deep inference formalism) polynomially simulate each other [18].

DEFINITION 5
We define a *section* of a derivation in the following way:

- Every atom a has one section, a.
- If A is a section of ϕ, and B is a section of ψ, then $A \star B$ is a section of $\phi \star \psi$, and QxA is a section of $Qx\phi$.
- If A is a section of $\begin{array}{c} B \\ \phi_1 \Vert \\ C \end{array}$ or $\begin{array}{c} D \\ \phi_2 \Vert \\ E \end{array}$ and $\phi = \begin{array}{c} B \\ \phi_1 \Vert \\ \rho \dfrac{C}{D} \\ \phi_2 \Vert \\ E \end{array}$ then A is a section of ϕ.

The premise and conclusion of a derivation are, respectively, the uppermost section and lowermost section of the derivation. A *proof* of A is a derivation with premise t and conclusion A, sometimes written $\begin{array}{c} \phi \Vert \\ A \end{array}$.

DEFINITION 6
We define the rewriting system Seq as containing the following two rewrites S_l and S_r:

$$
= \dfrac{\begin{array}{c} A \\ \Vert \\ K \left\{ \rho_1 \dfrac{A_1}{B_1} \right\} \{A_2\} \end{array}}{\begin{array}{c} K\{B_1\} \left\{ \rho_2 \dfrac{A_2}{B_2} \right\} \\ \Vert \\ B \end{array}} \quad \xleftarrow{S_l} \quad \begin{array}{c} A \\ \Vert \\ K \left\{ \rho_1 \dfrac{A_1}{B_1} \right\} \left\{ \rho_2 \dfrac{A_2}{B_2} \right\} \\ \Vert \\ B \end{array} \quad \xrightarrow{S_r} \quad = \dfrac{\begin{array}{c} A \\ \Vert \\ K\{A_1\} \left\{ \rho_2 \dfrac{A_2}{B_2} \right\} \end{array}}{\begin{array}{c} K \left\{ \rho_1 \dfrac{A_1}{B_1} \right\} \{B_2\} \\ \Vert \\ B \end{array}}
$$

If ϕ is in normal form w.r.t. Seq, we say ϕ is in *sequential form*. If $\phi \longrightarrow^*_{\mathsf{Seq}} \psi$ and ψ is in sequential form, we say that ψ is a *sequentialization* of ϕ.

PROPOSITION 7
A derivation ϕ is in sequential form iff. it is in the following form, where ρ_i are all the non-equality rules:

$$
= \frac{A}{\dfrac{K_1\left\{\rho_1 \dfrac{A_1}{B_1}\right\}}{= \dfrac{\vdots}{= \dfrac{K_n\left\{\rho_n \dfrac{A_n}{B_n}\right\}}{= \, B}}}}
$$

DEFINITION 8
A *closed* derivation is one where every section of the derivation is a sentence (i.e. a formula with no free variables), and is *regular* if no variable is used in two different quantifiers.

We define relative strength and equivalence between proof systems in the standard fashion.

DEFINITION 9
An inference rule ρ is *admissible* for a proof system S if for every proof $\overset{\phi \Vert \mathsf{S} \cup \{\rho\}}{A}$ there is a proof $\overset{\phi' \Vert \mathsf{S}}{A}$.

DEFINITION 10
An inference rule ρ is *derivable* for a proof system S if for every instance $\rho \frac{A}{B}$ of ρ there is a derivation $\begin{array}{c} A \\ \phi \Vert \mathsf{S} \\ B \end{array}$.

2.1 KS *and* SKS*,* KSq *and* SKSq

We define four basic open deduction systems for classical logic, two for propositional logic, two for classical; two cut-free systems and two with cut. These are standard in the literature [8, 10] and allow us to ground the material developed below in the existing deep inference literature.

DEFINITION 11
In Figure 3, we define the proof systems $\mathsf{KS}, \mathsf{SKS}, \mathsf{KSq}$ and SKSq.

Cut elimination for propositional logic is a standard deep inference theorem.

THEOREM 12
$\mathsf{ai}\uparrow$ is admissible for all systems $\mathsf{KS} \subseteq \mathsf{S} \subseteq \mathsf{SKS} \backslash \{\mathsf{ai}\uparrow\}$

PROOF. Proofs can be found in many places, e.g. Theorem 1.55 of [32]. □

SKS

$$\mathsf{ai}{\uparrow}\,\frac{a \wedge \bar{a}}{\mathsf{f}} \qquad \mathsf{ac}{\uparrow}\,\frac{a}{a \wedge a} \qquad \mathsf{aw}{\uparrow}\,\frac{a}{\mathsf{t}}$$

KS

$$\mathsf{ai}{\downarrow}\,\frac{\mathsf{t}}{a \vee \bar{a}} \qquad \mathsf{ac}{\downarrow}\,\frac{a \vee a}{a} \qquad \mathsf{aw}{\downarrow}\,\frac{\mathsf{f}}{a}$$

$$\mathsf{s}\,\frac{A \wedge (B \vee C)}{(A \wedge B) \vee C} \qquad \mathsf{m}\,\frac{(A \wedge B) \vee (C \wedge D)}{(A \vee C) \wedge (B \vee D)}$$

SKSq

$$\mathsf{u}{\uparrow}\,\frac{\exists x A \wedge \forall x B}{\exists x (A \wedge B)} \qquad \mathsf{m_2}{\uparrow}\,\frac{\exists x (A \wedge B)}{\exists x A \wedge \exists x B}$$

$$\mathsf{n}{\uparrow}\,\frac{\forall x A}{[\tau \Rightarrow x]A} \qquad \mathsf{m_1}{\uparrow}\,\frac{\forall x (A \wedge B)}{\forall x A \wedge \forall x B}$$

KSq

$$\mathsf{n}{\downarrow}\,\frac{[\tau \Rightarrow x]A}{\exists x A} \qquad \mathsf{m_1}{\downarrow}\,\frac{\exists x A \vee \exists x B}{\exists x [A \vee B]}$$

$$\mathsf{u}{\downarrow}\,\frac{\forall x [A \vee B]}{\forall x A \vee \exists x B} \qquad \mathsf{m_2}{\downarrow}\,\frac{\forall x A \vee \forall x B}{\forall x [A \vee B]}$$

$$\begin{array}{cc} A \wedge \mathsf{t} = A & A \vee \mathsf{f} = A \\ \mathsf{t} \vee \mathsf{t} = \mathsf{t} & \mathsf{f} \wedge \mathsf{f} = \mathsf{f} \\ A \wedge B = B \wedge A & A \vee B = B \vee A \\ A \wedge (B \wedge C) = (A \wedge B) \wedge C & A \vee [B \vee C] = [A \vee B].C \end{array}$$

$$\begin{array}{lll} \forall x A = \forall z([z \Rightarrow x]A) & \forall x \forall y A = \forall y \forall x A & \forall x \mathsf{t} = \mathsf{t} = \exists x \mathsf{t} \\ \exists z A = \exists z([z \Rightarrow x]A) & \exists x \exists y A = \exists y \exists x A & \forall x \mathsf{f} = \mathsf{f} = \exists x \mathsf{f} \end{array}$$

FIGURE 3. The four core classical proof systems. The two dimensions are cut-free/cut-full (S) and propositional/first-order (q). Below, in the middle box, is the equality relation for propositional logic, which is extended by the rules in the bottom box for first-order logic.

We also introduce the *rules of passage* (also known as *retract rules* [10]), originally described by Herbrand [20, 21]. Since they are essentially used in his work as rewriting rules, they can be considered deep inference rules *avant la lettre*.

DEFINITION 13
The following eight rules are the rules of passage

$$\mathsf{r1}{\downarrow}\,\frac{\forall x (A \vee B)}{\forall x A \vee B} \qquad \mathsf{r2}{\downarrow}\,\frac{\forall x (A \wedge B)}{\forall x A \wedge B} \qquad \mathsf{r3}{\downarrow}\,\frac{\exists x (A \vee B)}{\exists x A \vee B} \qquad \mathsf{r4}{\downarrow}\,\frac{\exists x (A \wedge B)}{\exists x A \wedge B}$$

$$\mathsf{r1}{\uparrow}\,\frac{\exists x A \wedge B}{\exists x (A \wedge B)} \qquad \mathsf{r2}{\uparrow}\,\frac{\exists x A \vee B}{\exists x (A \vee B)} \qquad \mathsf{r3}{\uparrow}\,\frac{\forall x A \wedge B}{\forall x (A \wedge B)} \qquad \mathsf{r4}{\uparrow}\,\frac{\forall x A \vee B}{\forall x (A \vee B)}$$

where $x \in FV(B)$. we refer to the down rules collectively as $\mathsf{RP}_{\downarrow} = \{\mathsf{r1}{\downarrow}, \mathsf{r2}{\downarrow}, \mathsf{r3}{\downarrow}, \mathsf{r4}{\downarrow}\}$, the up rules as $\mathsf{RP}_{\uparrow} = \{\mathsf{r1}{\uparrow}, \mathsf{r2}{\uparrow}, \mathsf{r3}{\uparrow}, \mathsf{r4}{\uparrow}\}$ and all the rules of passage as $\mathsf{RP} = \mathsf{RP}_{\downarrow} \cup \mathsf{RP}_{\uparrow}$.

The rules of passage allow for prenexification and deprenexification of formulae. Since there are formulae whose cut-free sequent calculus proofs are non-elementarily shorter than their prenexified forms [1, 7, 33], adding the four rules $\{\mathsf{r1}{\uparrow}, \mathsf{r2}{\uparrow}, \mathsf{r3}{\uparrow}, \mathsf{r4}{\uparrow}\}$ to a cut-free first-order system leads to significantly shorter proofs.

THEOREM 14
SKS and KS are sound and complete for classical propositional logic, SKSq and KSq are sound and complete for first order logic

3 Herbrand Proofs

As discussed in the introduction, we will present two different conceptions of representing the 'Herbrand content' of a proof: Herbrand proofs and expansion proofs. For each we define both a deep inference proof system—$\mathsf{KSh1}$ and $\mathsf{KSh2}$—and a class of proofs in each system that corresponds to Herbrand or expansion proofs, respectively. First, we will present $\mathsf{KSh1}$ and Herbrand proofs.

3.1 $\mathsf{KSh1}$ *and Herbrand proofs*

As discussed in the introduction, Herbrand proofs consist of the following four steps:

1. Expansion of existential subformulae.
2. Prenexification/elimination of universal quantifiers.
3. Term assignment.
4. Propositional tautology check.

In [10], it is shown that all four of these steps can be carried out by inference rules in a deep inference system. To do so, we need to define a contraction rule that only operates on existential formula.

DEFINITION 15
We define the rule $\mathsf{qc}{\downarrow}$ to restrict contraction just to existential formulae:

$$\mathsf{qc}{\downarrow}\frac{\exists x A \vee \exists x A}{\exists x A}$$

PROPOSITION 16
$\mathsf{c}\downarrow$ is derivable for $\{\mathsf{ac}{\downarrow}, \mathsf{m}, \mathsf{qc}{\downarrow}, \mathsf{m}_2{\downarrow}\}$. $\mathsf{qc}{\downarrow}$ is derivable for $\{\mathsf{ac}{\downarrow}, \mathsf{m}, \mathsf{m}_1{\downarrow}, \mathsf{m}_2{\downarrow}\}$.

PROOF. Straightforward. □

DEFINITION 17
We define a proof system for FOL, $\mathsf{KSh1}$:

$$\mathsf{KSh1} = \mathsf{KS} + \boxed{\begin{array}{ccc} \mathsf{r1}{\downarrow}\dfrac{\forall x[A \vee B]}{[\forall x A \vee B]} & \mathsf{r2}{\downarrow}\dfrac{\forall x(A \wedge B)}{(\forall x A \wedge B)} & \mathsf{n}{\downarrow}\dfrac{A\{x \Leftarrow t\}}{\exists x A} \\[2ex] \mathsf{r3}{\downarrow}\dfrac{\exists x[A \vee B]}{[\exists x A \vee B]} & \mathsf{r4}{\downarrow}\dfrac{\exists x(A \wedge B)}{(\exists x A \wedge B)} & \mathsf{qc}{\downarrow}\dfrac{\exists x A \vee \exists x A}{\exists x A} \end{array}}$$

with the usual equality relation for first-order logic.

Following [10], we define a Herbrand proof in the context of $\mathsf{KSh1}$ in the following way.

$$
\mathsf{qc}\!\downarrow\dfrac{\mathsf{r1}\!\downarrow\dfrac{\mathsf{r1}\!\downarrow\dfrac{=\dfrac{\mathsf{t}}{\forall y_1\forall y_2\boxed{=\dfrac{\mathsf{ai}\!\downarrow\dfrac{\mathsf{t}}{Py_1\vee\bar{P}y_1}\vee\left(\mathsf{aw}\!\downarrow\dfrac{\mathsf{f}}{\bar{P}c}\vee\mathsf{aw}\!\downarrow\dfrac{\mathsf{f}}{Py_2}\right)}{\left(\bar{P}c\vee Py_1\right)\vee\left(\bar{P}y_1\vee Py_2\right)}}}}{\forall y_1\boxed{\left(\bar{P}c\vee Py_1\right)\vee\mathsf{n}\!\downarrow\dfrac{\forall y_2\left(\bar{P}y_1\vee Py_2\right)}{\exists x_2\forall y_2\left(\bar{P}x_2\vee Py_2\right)}}}}{\mathsf{n}\!\downarrow\dfrac{\forall y_1\left(\bar{P}c\vee Py_1\right)}{\exists x_1\forall y_1\left(\bar{P}x_1\vee Py_1\right)}\vee\exists x_2\forall y_2\left(\bar{P}x_2\vee Py_2\right)}}{\exists x\forall y\left(\bar{P}x\vee Py\right)}
$$

FIGURE 4. A $\mathsf{KSh1}$ proof of the 'drinker's formula'.

DEFINITION 18
A closed $\mathsf{KSh1}$ proof is a *Herbrand proof* if it is in the following form:

$$
\begin{array}{c}
\Vert\,\mathsf{KS} \\
\forall\vec{x}B\{\vec{y}\Leftarrow\vec{t}\} \\
\Vert\,\{\mathsf{n}\!\downarrow\} \\
Q\{B\} \\
\Vert\,\mathsf{RP}_{\downarrow} \\
A' \\
\Vert\,\{\mathsf{qc}\!\downarrow\} \\
A
\end{array}
$$

where $Q\{\ \}$ is a context consisting only of quantifiers and B is quantifier-free.

THEOREM 19 (Herbrand's theorem for cut-free proofs).
Let $\begin{smallmatrix}\phi\,\Vert\,\mathsf{KSq}\\ A\end{smallmatrix}$. Then we can construct a Herbrand proof of A.

REMARK 20 This proposition and proof are essentially the same the proof of Herbrand's theorem from a cut-free system in [10, Theorem 4.2]. However, the proof is worth reworking in the open deduction formalism, and not simply as a corollary to cut elimination.

Before proving the theorem, we state and prove an important lemma, which requires an omitted proposition that can be found proven in [32].

PROPOSITION 21
$\mathsf{n}\!\uparrow$ is admissible for KSq

PROOF. A proof of an equivalent statement can be found as Proposition 3.22 in [32]. □

$$
\mathsf{qc}{\downarrow}\,\dfrac{\mathsf{r3}{\downarrow}\,\dfrac{\mathsf{n}{\downarrow}\,\dfrac{=\,\dfrac{\mathsf{t}}{\forall y_1\forall y_2\boxed{=\,\dfrac{\mathsf{ai}{\downarrow}\dfrac{\mathsf{t}}{Py_1\vee\bar{P}y_1}\vee\left(\mathsf{aw}{\downarrow}\dfrac{\mathsf{f}}{\bar{P}c}\vee\mathsf{aw}{\downarrow}\dfrac{\mathsf{f}}{Py_2}\right)}{\left(\bar{P}c\vee Py_1\right)\vee\left(\bar{P}y_1\vee Py_2\right)}}}}{\exists x_1\boxed{\mathsf{r1}{\downarrow}\,\dfrac{\forall y_1\boxed{\mathsf{r3}{\downarrow}\,\dfrac{\mathsf{n}{\downarrow}\,\dfrac{\left(\bar{P}x_1\vee Py_1\right)\vee\left(\bar{P}y_1\vee Py_2\right)}{\exists x_2\boxed{\mathsf{r1}{\downarrow}\,\dfrac{\forall y_2\left(\left(\bar{P}x_1\vee Py_1\right)\vee\left(\bar{P}x_2\vee Py_2\right)\right)}{\left(\bar{P}x_1\vee Py_1\right)\vee\forall y_2\left(\bar{P}x_2\vee Py_2\right)}}}}{\left(\bar{P}x_1\vee Py_1\right)\vee\exists x_2\forall y_2\left(\bar{P}x_2\vee Py_2\right)}}}{\forall y_1\left(\bar{P}x_1\vee Py_1\right)\vee\exists x_2\forall y_2\left(\bar{P}x_2\vee Py_2\right)}}}}{\exists x_1\forall y_1\left(\bar{P}x_1\vee Py_1\right)\vee\exists x_2\forall y_2\left(\bar{P}x_2\vee Py_2\right)}}{\exists x\forall y\left(\bar{P}x\vee Py\right)}
$$

FIGURE 5. A Herbrand proof of the drinker's formula.

LEMMA 22
We can carry out the following proof transformation:

$$
\begin{array}{c} \phi \,\Vert\, \mathsf{KSq}\setminus\{\mathsf{m}_1{\downarrow}\} \\ A \end{array}
\quad\longrightarrow\quad
\begin{array}{c} \Vert\, \mathsf{KS}\cup\{\mathsf{n}{\downarrow}\} \\ Q\{A_P\} \\ \Vert\, \mathsf{RP}_{\downarrow} \\ A \end{array}
$$

where $Q\{\ \}$ is a sequence of quantifiers and A_P is the formula obtained by removing all quantifiers from A, and $\mathsf{RP}_{\downarrow} = \{\mathsf{r1}{\downarrow}, \mathsf{r2}{\downarrow}, \mathsf{r3}{\downarrow}, \mathsf{r4}{\downarrow}\}$.

PROOF. We proceed by induction on the length of ϕ, with the base case being trivial. If the final rule in ϕ is in KS, the inductive step is also trivial. We use the rewrite

$$
\mathsf{u}{\downarrow}\,\dfrac{\forall x(A \vee B)}{\forall xA \vee \exists xB}
\quad\longrightarrow\quad
\mathsf{r1}{\downarrow}\,\dfrac{\forall x\boxed{A \vee \mathsf{n}{\downarrow}\,\dfrac{B}{\exists xB}}}{\forall xA \vee \exists xB}
$$

to replace $\mathsf{u}{\downarrow}$ with $\{\mathsf{r1}{\downarrow}, \mathsf{n}{\downarrow}\}$, with $\mathsf{r1}{\downarrow}$ simply getting absorbed into $\mathsf{RP}_{\downarrow}$. Thus, we are left with just $\mathsf{m}_2{\downarrow}$ and $\mathsf{n}{\downarrow}$ to deal with.

We can eliminate instances of $\mathsf{m}_2{\downarrow}$ in the following way:

$$
\begin{array}{c} \phi \,\| \, \mathsf{KSq}\backslash\{\mathsf{m}_1{\downarrow}\} \\ K\left\{ \mathsf{m}_2{\downarrow}\,\dfrac{\forall xA \vee \forall y[y\Rightarrow x]B}{\forall x(A\vee B)} \right\} \end{array}
\xrightarrow{IH}
\begin{array}{c} \phi' \,\|\, \mathsf{KS}\cup\{\mathsf{n}{\downarrow}\} \\ Q_1\{\forall x Q_2\{\forall y Q_3\{A_P \vee [y\Rightarrow x]B_P\}\}\} \\ \|\,\mathsf{RP}_{\downarrow} \\ K\left\{ \mathsf{m}_2{\downarrow}\,\dfrac{\forall xA \vee \forall y[y\Rightarrow x]B}{\forall x(A\vee B)} \right\} \end{array}
\longrightarrow
$$

$$
\begin{array}{c} \phi' \,\|\, \mathsf{KS}\cup\{\mathsf{n}{\downarrow}\} \\ Q_1\left\{\forall x Q_2\left\{ \mathsf{n}{\uparrow}\,\dfrac{\forall y Q_3\{A_P \vee [y\Rightarrow x]B_P\}}{Q_3\{A_P\vee B_P\}} \right\}\right\} \\ \|\,\mathsf{RP}_{\downarrow} \\ K\{\forall x(A\vee B)\} \end{array}
\xrightarrow{\text{Prop. 21}}
\begin{array}{c} \phi' \,\|\, \mathsf{KS}\cup\{\mathsf{n}{\downarrow}\} \\ Q_1\{\forall x Q_2\{Q_3\{A_P\vee B_P\}\}\} \\ \|\,\mathsf{RP}_{\downarrow} \\ K\{\forall x(A\vee B)\} \end{array}
$$

Instances of $\mathsf{n}{\downarrow}$ are permuted above the rules of passage:

$$
\begin{array}{c} \phi \,\|\, \mathsf{KSq}\backslash\{\mathsf{m}_1{\downarrow}\} \\ K\left\{ \mathsf{n}{\downarrow}\,\dfrac{[t\Rightarrow x]A}{\exists xA} \right\} \end{array}
\xrightarrow{IH}
\begin{array}{c} \phi \,\|\, \mathsf{KS}\cup\{\mathsf{n}{\downarrow}\} \\ Q_1\{Q_2\{[t\Rightarrow x]A_P\}\} \\ \|\,\mathsf{RP}_{\downarrow} \\ K\left\{ \mathsf{n}{\downarrow}\,\dfrac{[t\Rightarrow x]A}{\exists xA} \right\} \end{array}
\longrightarrow
\begin{array}{c} \phi \,\|\, \mathsf{KS}\cup\{\mathsf{n}{\downarrow}\} \\ Q_1\left\{ \mathsf{n}{\downarrow}\,\dfrac{Q_2\{[t\Rightarrow x]A_P\}}{\exists x Q_2\{[t\Rightarrow x]A_P\}} \right\} \\ \|\,\mathsf{RP}_{\downarrow} \\ K\{\exists xA\} \end{array}
$$

□

PROOF OF THEOREM 19. We work in stages, creating one section of the Herbrand proof at a time.

1. First, we refactorize contraction as $\{\mathsf{ac}{\downarrow}, \mathsf{m}, \mathsf{qc}{\downarrow}, \mathsf{m}_2{\downarrow}\}$ instead of $\{\mathsf{ac}{\downarrow}, \mathsf{m}, \mathsf{m}_1{\downarrow}, \mathsf{m}_2{\downarrow}\}$. We then use the following rewrites to permute $\mathsf{qc}{\downarrow}$ down:

$$
\mathsf{qc}{\downarrow}-\rho_1: \quad \rho\,\dfrac{K\left\{ \mathsf{qc}{\downarrow}\,\dfrac{\exists xA \vee \exists xA}{\exists xA} \right\}}{K'\{\exists xA\}} \longrightarrow \rho\,\dfrac{K\{\exists xA\vee\exists xA\}}{K'\left\{ \mathsf{qc}{\downarrow}\,\dfrac{\exists xA \vee \exists xA}{\exists xA} \right\}}
$$

$$
\mathsf{qc}{\downarrow}-\rho_2: \quad \mathsf{qc}{\downarrow}\,\dfrac{\exists xA \vee \exists xA}{\exists x \boxed{\rho\,\dfrac{A}{B}}} \longrightarrow \mathsf{qc}{\downarrow}\,\dfrac{\exists x\boxed{\rho\,\dfrac{A}{B}} \vee \exists x\boxed{\rho\,\dfrac{A}{B}}}{\exists xB}
$$

   Each of these reductions, if applied to the bottommost instance of $\mathsf{qc}{\downarrow}$, reduces the number of rules below the bottommost instance of $\mathsf{qc}{\downarrow}$.

2. By Lemma 22 we can now separate the remaining proof into a top half with $\mathsf{KS}\cup\{\mathsf{n}{\downarrow}\}$, and a bottom half consisting of $\mathsf{RP}_{\downarrow}$.
3. Since every other first-order rule is now eliminated from the proof, it is straightforward to permute $\mathsf{n}{\downarrow}$ rules down the proof.

REMARK 23 The proof of Theorem 19 works exactly the same in the if KS is supplemented by an atomic cut rule, $\mathsf{ai}\uparrow$. Therefore, as Brünnler notes, only cuts with quantified eigenformulae need be eliminated to prove Herbrand's theorem.

□

PROPOSITION 24
Given a Herbrand proof ϕ of A, we can construct a KSq proof of A.

PROOF. Omitted, can be found as Proposition 4.9 of [32]. □

4 Expansion Proofs

4.1 Introduction

In [30], Miller generalizes the concept of the Herbrand expansion to higher order logic, representing the witness information in a tree structure, and explicit transformations between these 'expansion proofs' and cut-free sequent proofs are provided. Miller's presentation of expansion proofs lacked some of the usual features of a formal proof system, crucially composition by an eliminable cut, but extensions in this direction have been carried out by multiple authors. In [19], Heijltjes presents a system of 'proof forests', a graphical formalism of expansion proofs with cut and a local rewrite relation that performs cut elimination. Similar work has been carried out by McKinley [29] and more recently by Hetzl and Weller [6] and Alcolei et al. [2].

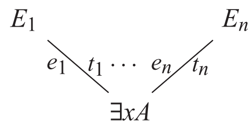
4.2 Expansion trees

REMARK 25 In this section, we will frequently use $\star$ in place of $\wedge$ and $\vee$, and Q in place of $\forall$ and $\exists$ if both cases can be combined into one. For clarity, we will sometimes distinguish between connectives in expansion trees, $\star_E$, and in formulae/derivations, $\star_F$.

DEFINITION 26
We define *expansion trees*, the two functions Sh (shallow) and Dp (deep) from expansion trees to formulae, a set of *eigenvariables* $EV(E)$ for each expansion tree, and a partial function Lab from edges to terms, following [30], [19] and [13]:

1. Every literal A (including the units t and f) is an expansion tree. $Sh(A) := A$, $Dp(A) := A$, and $EV(A) = \emptyset$.
2. If E_1 and E_2 are expansion trees with $EV(E_1) \cap EV(E_2) = \emptyset$, then $E_1 \star E_2$ is an expansion tree, with $Sh(E_1 \star_E E_2) := Sh(E_1) \star_F Sh(E_2)$, $Dp(E_1 \star_E E_2) := Dp(E_1) \star_F Dp(E_2)$, and $EV(E_1 \star E_2) = EV(E_1) \cup EV(E_2)$. We call $\star$ a $\star$-node and each unlabelled edge e_i connecting the $\star$-node to E_i a $\star$-edge. We represent $E_1 \star E_2$ as:
$$\begin{array}{ccc} E_1 & & E_2 \\ e_1\diagdown & & \diagup e_2 \\ & \star & \end{array}$$
3. If E' is an expansion tree s.t. $Sh(E') = A$ and $x \notin EV(E')$, then $E = \forall xA +^x E'$ is an expansion tree with $Sh(E) := \forall xA$, $Dp(E) := Dp(E')$, and $EV(E) := EV(E') \cup \{x\}$. We call $\forall xA$ a $\forall$-node and the edge e connecting the $\forall$-node and E' a $\forall$-edge, with $Lab(e) = x$. We represent E as:
$$\begin{array}{c} E' \\ e \,\big|\, x \\ \forall xA \end{array}$$

4. If $t_1, \ldots, t_n$ are terms ($n \geq 0$), and $E_1, \ldots, E_n$ are expansion trees s.t. $x \notin EV(E_i)$ and $EV(E_i) \cap EV(E_j) = \emptyset$ for all $1 \leq i < j \leq n$, and $Sh(E_i) = [t_i \Rightarrow x]A$, then $E = \exists x A +^{t_1} E_1 +^{t_2} \cdots +^{t_n} E_n$ is an expansion tree, where $Sh(E) := \exists x A$, $Dp(E) := Dp(E_1) \vee \cdots \vee Dp(E_n)$ (with $Dp(E) = \mathsf{f}$ if $n = 0$), and $EV(E) = \bigcup_1^n EV(E_i)$. We call $\exists x A$ an $\exists$-node and each edge e_i connecting the $\exists$-node with E_i an $\exists$-edge, with $Lab(e_i) = t_i$. We represent E as:



REMARK 27 Let ρ be a permutation of $[1 \ldots n]$. We consider the expansion trees $\exists x A +^{t_1} E_1 +^{t_2} \cdots +^{t_n} E_n$ and $\exists x A +^{t_{\rho(1)}} E_{\rho(1)} \cdots +^{t_{\rho(n)}} E_{\rho(n)}$ equal. Our trees are also presented the other way up to usual, e.g. [19]. This is so that they are the same way up as the deep inference proofs we will translate them to below.

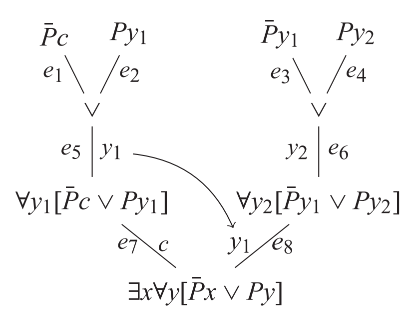
DEFINITION 28
Let E be an expansion tree and let $<_E^-$ be the relation on the edges in E defined by:

- $e <_E^- e'$ if the node directly below e is the node directly above e'.
- $e <_E^- e'$ if e is an $\exists$-edge with $Lab(e) = t$, $x \in FV(t)$, e' is a $\forall$-edge and $Lab(e') = x$. In this case, we say e' *points to* e.

The *dependency relation* of E, $<_E$, is the transitive closure of $<_E^-$.

DEFINITION 29
An expansion tree E is *correct* if $<_E$ is acyclic and $Dp(E)$ is a tautology. We can then call E an *expansion proof* of $Sh(E)$.

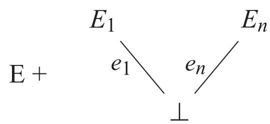
EXAMPLE 30
Below is an expansion tree E, with $Sh(E) = \exists x \forall y(\bar{P}x \vee Py)$ and $Dp(E) = (\bar{P}c \vee Py_1) \vee (\bar{P}y_1 \vee Py_2)$. The tree is presented with all edges explicitly named, to define the dependency relation below, as well as the labels for the $\exists$-edges and $\forall$-edges.



The dependency relation is generated by the following inequalities: $e_3, e_4 < e_6 < e_8$ and $e_1, e_2 < e_5 < e_7$ and $e_8 < e_5$. e_5 points to e_8. As this dependency relation is acyclic and $(\bar{P}c \vee Py_1) \vee (\bar{P}y_1 \vee Py_2)$ is a tautology, E is correct, and thus an expansion proof.

4.3 Expansion proofs with cut

DEFINITION 31
If $E, E_1, \ldots, E_n$ are expansion trees, with $Sh(E_i) = \exists x A_i \wedge \forall x - A_i$ for each E_i, then $EC = E + \bot(E_1, \ldots, E_n)$ is an *expansion tree with cut*. We call $\bot$ the *cut node*, and each edge e_i connecting $\bot$ and E_i a *cut edge*. If $FV(\exists x A_i) = \emptyset$, then we say e_i is a *closed* cut edge. We represent EC as:



We extend the deep and shallow functions to expansion trees with cuts: $Sh(E + \bot(E_1, \ldots, E_n)) = A$, $Dp(E + \bot(E_1, \ldots, E_n)) = Dp(E) \vee Dp(E_1) \vee \cdots \vee Dp(E_n)$.

Extending the notion of correctness to expansion trees with cut is straightforward.

DEFINITION 32
The dependency relation for expansion trees with cut is the same as that for expansion trees, with the following addition to the definition of $<_E^-$:

- $e <_E^- e'$ if e is a cut edge connecting $\bot$ and E_i, $Sh(E_i) = \forall y A_i . \exists y \bar{A}_i$, $x \in FV(A_i)$, e' is a $\forall$-edge and $Lab(e') = x$. We still say that e' *points to* e.

The correctness criteria for expansion trees with cuts are the same for expansion trees, giving us *expansion proofs with cut*.

REMARK 33 If every cut edge is closed for an expansion tree with cut, then the correctness criteria is exactly the same as for the expansion tree obtained by replacing the cut node with a series of $\wedge$ nodes.

At this point, we are close to being able to borrow the cut elimination method from Heijltjes' *Proof Forests* formalism [19]. However, proof forests are only a subclass of what we define as expansion trees here. Therefore, it will be useful to properly define this subclass, as well as the class expansion proof with closed cuts, which will be another useful subclass later on.

DEFINITION 34
A *prenex* expansion tree is an expansion tree where no Q-node is above a $\star$-node.

If $E = E_1 \vee \cdots \vee E_n]$ with E_i prenex expansion trees, then E is a *forest-style expansion tree*. If E is correct it is a *forest-style expansion proof*.

If $E = E' + \bot(E_1 \wedge F_1, \ldots, E_n \wedge F_n)$ with E' a forest-style expansion tree, and $E_1, F_1, \ldots, E_n, F_n$ are prenex expansion trees, then E is a *forest-style expansion tree with cut*. If F_c is correct, then it is a *forest-style expansion proof with cut*.

If E is an expansion tree with cut with every cut edge closed, we say that E is a *expansion tree with closed cut*. If F_c is correct, then it is a *expansion proof with closed cut*.

CONVENTION 35
We consider every expansion tree/proof to also be an expansion tree/proof with cut.

THEOREM 36
If there is a forest-style expansion proof with cut F_c with $Sh(F_c) = A$, then we can construct from it a cut-free expansion proof $E_F = E_1 \vee \cdots \vee E_n$ where E_i are prenex expansion trees and $Sh(E_F) = A$.

PROOF. [19, Proposition 16 and Theorem 21] □

4.4 KSh2 *and Herbrand normal form*

To aid the translation between open deduction proofs and expansion proofs, we introduce a slightly different proof system to KSh1. It involves two new rules.

DEFINITION 37
We define the rule h↓, which we call a *Herbrand expander* and the rule ∃w, which we call *existential weakening*:

$$\mathsf{h}\!\downarrow \frac{\exists x A \vee [t \Rightarrow x]A}{\exists x A} \qquad \exists\mathsf{w}\!\downarrow \frac{\mathsf{f}}{\exists x A}$$

For technical reasons,, we insist that $[t \Rightarrow x]A$ is in fact $[t \Rightarrow x]A'$, where A' is an α-equivalent formula to A with fresh variables for all quantifiers, but for simplicity we will usually denote it A.

REMARK 38 Unlike the n↓ rule, the h↓ Herbrand expander rule is invertible. Similar rules have been used in first-order sequent calculus systems for automated reasoning, such as Kanger's LC [14, 26] and also in sequent systems for translation to expansion proofs [6].

DEFINITION 39
We define the first-order proof system:

$$\mathsf{KSh2} = \mathsf{KS} + \boxed{\begin{array}{cc} \mathsf{r1}\!\downarrow \dfrac{\forall x[A \vee B]}{[\forall x A \vee B]} & \mathsf{h}\!\downarrow \dfrac{\exists x A \vee [t \Rightarrow x]A}{\exists x A} \\ \mathsf{r2}\!\downarrow \dfrac{\forall x(A \wedge B)}{(\forall x A \wedge B)} & \exists\mathsf{w}\!\downarrow \dfrac{\mathsf{f}}{\exists x A} \end{array}}$$

This system uses the usual equality relation for first-order logic.

REMARK 40 The ∃w ↓ rule is derivable for KSh2\{∃w ↓}, but we explicitly include it so that we can restrict weakening instances in certain parts of proofs.

DEFINITION 41
If ϕ is a closed KSh2 proof in the following form, where $\forall\vec{x}$ is a list of universal quantifiers with distinct variables, and $Lo(\phi)$ is regular and in sequential form, we say ϕ is in *Herbrand Normal Form* (HNF):

$$\begin{array}{c} Up((\phi)) \, \Big\Vert \, \mathsf{KS} \\ \forall\vec{x} H_\phi((A)) \\ \Big\Vert \, \{\exists\mathsf{w}\!\downarrow\} \\ \forall\vec{x} H^+_\phi((A)) \\ Lo(\phi) \, \Big\Vert \, \{\mathsf{r1}\!\downarrow, \mathsf{r2}\!\downarrow, \mathsf{h}\!\downarrow\} \\ A \end{array}$$

$$
\mathsf{r1}\!\downarrow \dfrac{
=\dfrac{=\dfrac{\mathsf{t}}{\forall y_1 \forall y_2 \boxed{\mathsf{ai}\!\downarrow \dfrac{\mathsf{t}}{Py_1 \vee \bar{P}y_1} \vee \left(\mathsf{aw}\!\downarrow \dfrac{\mathsf{f}}{\bar{P}c} \vee \mathsf{aw}\!\downarrow \dfrac{\mathsf{f}}{Py_2}\right)}}}{\forall y_1 \boxed{\mathsf{r1}\!\downarrow \dfrac{\forall y_2 \boxed{\left(\mathsf{w}\!\downarrow \dfrac{\mathsf{f}}{\exists x \forall y \left(\bar{P}x \vee Py\right)} \vee \left(\bar{P}y_1 \vee Py_2\right)\right)} \vee \left(\bar{P}c \vee Py_1\right)}{\mathsf{h}\!\downarrow \dfrac{\mathsf{r1}\!\downarrow \dfrac{\forall y_2 \left(\exists x \forall y \left(\bar{P}x \vee Py\right) \vee \left(\bar{P}y_1 \vee Py_2\right)\right)}{\exists x \forall y \left(\bar{P}x \vee Py\right) \vee \forall y_2 \left(\bar{P}y_1 \vee Py_2\right)}}{\exists x \forall y \left(\bar{P}x \vee Py\right)} \vee \left(\bar{P}c \vee Py_1\right)}}}
}{
\mathsf{h}\!\downarrow \dfrac{\exists x \forall y \left(\bar{P}x \vee Py\right) \vee \forall y_1 \left(\bar{P}c \vee Py_1\right)}{\exists x \forall y \left(\bar{P}x \vee Py\right)}
}
$$

FIGURE 6. A proof of the drinker's formula in HNF.

$H_\phi(A)$, the *Herbrand disjunction of A according to* ϕ, or just the *Herbrand disjunction of A*, contains no quantifiers, whereas $H^+_\phi(A)$, the *expansive Herbrand disjunction of A according to* ϕ, may contain quantifiers. $Up(\phi)$ is called the *upper part* of ϕ, and $Lo(\phi)$ the *lower part* of ϕ.

We also define proofs in HNF with cut. Notice that the cuts do not necessarily need to be at the bottom of the proof, and need not be closed.

DEFINITION 42
We define $\mathsf{KSh2c} = \mathsf{KSh2} \cup \{\mathsf{qi}\!\uparrow\}$, with $\mathsf{qi}\!\uparrow$ defined to be a cut that only operates on quantified formulae:

$$
\mathsf{qi}\!\uparrow \frac{\exists x A \wedge \forall x \bar{A}}{\mathsf{f}}
$$

DEFINITION 43
If ϕ is a closed $\mathsf{KSh2c}$ proof in the following form, then we say ϕ is in *Herbrand Normal Form with Cut* (HNFC)

$$
\begin{array}{c}
Up((\phi)) \, \prod \mathsf{KS} \\
\forall \vec{x} H_\phi((A)) \\
\Vert \, \{\exists \mathsf{w}\!\downarrow\} \\
\forall \vec{x} H^+_\phi((A)) \\
Lo(\phi) \, \Vert \, \{\mathsf{r1}\!\downarrow, \mathsf{r2}\!\downarrow, \mathsf{h}\!\downarrow, \mathsf{qi}\!\uparrow\} \\
A
\end{array}
$$

In this paper, we will only ever translate between proofs in HNF and Herbrand proofs when they are cut-free or with closed cuts. Therefore, a cut free translation between the two forms suffices.

PROPOSITION 44
A formula A has a proof in HNF iff. it has a Herbrand proof.

PROOF. Let ϕ be a proof of A in HNF. As $H_\phi(A)$ is the Herbrand expansion of A, it is straightforward to construct a Herbrand proof for A: one can infer the necessary $\mathsf{n}\downarrow$ and $\mathsf{qc}\downarrow$ rules by comparing $H_\phi(A)$ and A. Now let ϕ be a Herbrand proof. The order of the quantifiers in $Q\{\ \}$ (as in Definition 18) is used to build the HNF proof. Thus, we proceed by induction on the number of quantifiers in $Q\{\ \}$. If there are none, it is obviously trivial. We split the inductive step into two cases.

First, consider ϕ_1 of the form shown, where P is a quantifier-free context and $Q\{\ \} = \forall z Q'\{\ \}$. Clearly ϕ_2 is also a Herbrand proof, so by the IH the proof ϕ_3 in HNF is constructible, from which we can construct ϕ_4.

$$
\begin{array}{c}
\Vert\,\mathsf{KS} \\
\forall z\forall\vec{x}B\{\vec{y}\Leftarrow\vec{t}\} \\
\Vert\,\{\mathsf{n}\downarrow\} \\
\forall z Q'\{B\} \\
\Vert\,\{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow,\mathsf{r3}\downarrow,\mathsf{r4}\downarrow\} \\
P\{\forall z C'\} \\
\Vert\,\{\mathsf{qc}\downarrow\} \\
P\{\forall z C\} \\[1em]
\phi_1
\end{array}
\qquad
\begin{array}{c}
\Vert\,\mathsf{KS} \\
\forall\vec{x}B\{\vec{y}\Leftarrow\vec{t}\} \\
\Vert\,\{\mathsf{n}\downarrow\} \\
Q'\{B\} \\
\Vert\,\{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow,\mathsf{r3}\downarrow,\mathsf{r4}\downarrow\} \\
P\{C'\} \\
\Vert\,\{\mathsf{qc}\downarrow\} \\
P\{C\} \\[1em]
\phi_2
\end{array}
\qquad
\begin{array}{c}
Up(\phi_3)\,\Vert\,\mathsf{KS} \\
\forall\vec{x}H_{\phi_3}(P\{C\}) \\
\Vert\,\{\exists\mathsf{w}\downarrow\} \\
\forall\vec{x}H^+_{\phi_3}(P\{C\}) \\
Lo(\phi_3)\,\Vert\,\{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow,\mathsf{h}\downarrow\} \\
P\{C\} \\[1em]
\phi_3
\end{array}
\qquad
\begin{array}{c}
\forall z Up(\phi_3)\,\Vert\,\mathsf{KS} \\
\forall z\forall\vec{x}H_{\phi_3}(P\{C\}) \\
\Vert\,\{\exists\mathsf{w}\downarrow\} \\
\forall z\forall\vec{x}H^+_{\phi_3}(P\{C\}) \\
\forall z Lo(\phi_3)\,\Vert\,\{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow,\mathsf{h}\downarrow\} \\
\forall z P\{C\} \\
\Vert\,\{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow\} \\
P\{\forall z C\} \\[1em]
\phi_4
\end{array}
$$

In the same way, we consider the case where $Q\{\ \} = \exists z Q'\{\ \}$. Below we only show the case where there is no contraction acting on $\exists zC$, but the case with such a contraction is similar.

$$
\begin{array}{c}
\Vert\,\mathsf{KS} \\
\forall\vec{x}B\{\vec{y}\Leftarrow\vec{t}\}\{z\Leftarrow t\} \\
\Vert\,\{\mathsf{n}\downarrow\} \\
\exists z Q'\{B\} \\
\Vert\,\{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow,\mathsf{r3}\downarrow,\mathsf{r4}\downarrow\} \\
P\{\exists z C'\} \\
\Vert\,\{\mathsf{qc}\downarrow\} \\
P\{\exists z C\} \\[1em]
\phi_1
\end{array}
\qquad
\begin{array}{c}
\Vert\,\mathsf{KS} \\
\forall\vec{x}B\{\vec{y}\Leftarrow\vec{t}\}\{z\Leftarrow t\} \\
\Vert\,\{\mathsf{n}\downarrow\} \\
Q'\{B\}\{z\Leftarrow t\} \\
\Vert\,\{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow,\mathsf{r3}\downarrow,\mathsf{r4}\downarrow\} \\
P\{C'\{z\Leftarrow t\}\} \\
\Vert\,\{\mathsf{qc}\downarrow\} \\
P\{C\{z\Leftarrow t\}\} \\[1em]
\phi_2
\end{array}
\qquad
\begin{array}{c}
Up(\phi_3)\,\Vert\,\mathsf{KS} \\
\forall\vec{x}P\{D\{z\Leftarrow t\}\} \\
\Vert\,\{\exists\mathsf{w}\downarrow\} \\
\forall\vec{x}P\{D^+\{z\Leftarrow t\}\} \\
Lo(\phi_3)\,\Vert\,\{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow,\mathsf{h}\downarrow\} \\
P\{C\{z\Leftarrow t\}\} \\[1em]
\phi_3
\end{array}
\qquad
\begin{array}{c}
Up(\phi_3)\,\Vert\,\mathsf{KS} \\
\forall\vec{x}P\{D\{z\Leftarrow t\}\} \\
\Vert\,\{\exists\mathsf{w}\downarrow\} \\
\forall\vec{x}P\{\exists zC\vee D^+\{z\Leftarrow t\}\} \\
Lo(\phi_3)\,\Vert\,\{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow,\mathsf{h}\downarrow\} \\
P\left\{\mathsf{h}\downarrow\dfrac{\exists zC\vee C\{z\Leftarrow t\}}{\exists zC}\right\} \\[1em]
\phi_4
\end{array}
$$

where $P\{D\{z\Leftarrow t\}\} = H_{\phi_3}(P\{C\{z\Leftarrow t\}\})$ and $P\{D^+\{z\Leftarrow t\}\} = H^+_{\phi_3}(P\{C\{z\Leftarrow t\}\})$. $\square$

5 Translations Between Proofs in HNF and Expansion Proofs

Above, we gave translations between Herbrand proofs in $\mathsf{KSh1}$ and $\mathsf{KSh2}$ proofs in HNF. We will now give a translations between $\mathsf{KSh2c}$ proofs in HNFC and expansion proofs with cut, thus giving us a link between deep inference Herbrand proofs and expansion proofs. In the paper [31], we showed translations between Herbrand proofs and cut-free expansion proofs. Here, we will show translations between Herbrand proofs with cut and expansion proofs with cut, but these will be conservative extensions of the cut-free translations, so we use the same terminology.

REMARK 45 We extend the notion and syntax of contexts from derivations to expansion trees with cut. For the notion to make sense, a context can only take expansion trees with the same shallow formula.

5.1 HNF to expansion proofs

Before stating and proving the main theorem, we will define the map π_1 from KSh2c proofs to expansion proofs with cut (from now on in this section we will often omit 'with cut' if unambiguous), and then prove some lemmas to help prove that the dependency relation in all expansion proofs in the range of π_1 is acyclic.

DEFINITION 46
We define a map π_1' from the lower part of KSh2c proofs in HNFC to expansion trees in the following way, working from the bottom

On the conclusion of ϕ, we define π_1' as follows:

- $\pi_1'(B \star C) = \pi_1'(B) \star \pi_1'(C)$
- $\pi_1'(\forall xB) = \forall xB +^x \pi_1'(B)$
- $\pi_1'(\exists xB) = \exists xB$

The r1↓ and r2↓ rules are ignored by expansion trees, each h↓ rule adds a branch to a ∃-node, and each qi↓ rule adds another cut edge:

- If $\phi = \begin{array}{c} K\left\{ \mathsf{r1}{\downarrow}\dfrac{\forall x(B \vee C)}{\forall xB \vee C} \right\} \\ \phi' \| \\ A \end{array}$ then $\pi_1'(\phi) = \pi_1'\left(\begin{array}{c} K\{\forall xB \vee C\} \\ \phi' \| \\ A \end{array} \right)$.
- If $\phi = \begin{array}{c} K\left\{ \mathsf{r2}{\downarrow}\dfrac{\forall x(B \wedge C)}{\forall xB \wedge C} \right\} \\ \phi' \| \\ A \end{array}$ then $\pi_1'(\phi) = \pi_1'\left(\begin{array}{c} K\{\forall xB \wedge C\} \\ \phi' \| \\ A \end{array} \right)$.
- If $\pi_1'\left(\begin{array}{c} K\{\exists xB\} \\ \phi \| \\ A \end{array} \right) = K_{\pi_1}(\exists xB +^{\tau_1} E_1 + \cdots +^{\tau_n} E_n)$, then:

$$\pi_1'\left(\begin{array}{c} K\left\{ \mathsf{h}{\downarrow}\dfrac{\exists xB \vee [\tau_{n+1} \Rightarrow x]B}{\exists xB} \right\} \\ \phi \| \\ A \end{array} \right) = K_{\pi_1}(\exists xB +^{\tau_1} E_1 + \cdots +^{\tau_{n+1}} E_{n+1})$$

  where $E_{n+1} = \pi_1'([\tau_{n+1} \Rightarrow x]B)$.
- $\pi_1'\left(\begin{array}{c} K\{\mathsf{f}\} \\ \phi \| \\ A \end{array} \right) = E + \bot(E_1, \ldots, E_n)$, then:

$$\pi_1'\left(\begin{array}{c} K\left\{ \mathsf{qi}{\uparrow}\dfrac{\exists xB \wedge \forall xB}{\mathsf{f}} \right\} \\ \phi \| \\ A \end{array} \right) = E + \bot(E_1, \ldots, E_n, \pi_1'(\exists xB \wedge \forall xB))$$

We then define the map π_1 from KSh2 proofs in HNF to expansion trees as $\pi_1(\phi) = \pi_1'(Lo(\phi))$.

To show that $\pi_1(\phi)$ is an expansion proof, we need to prove that $\forall\vec{x}H_\phi(A)$ is a tautology and $<_E$ is acyclic. As $\forall\vec{x}H_\phi(A)$ has a proof in KS it is a tautology. Thus, all that is needed is the acyclicity of $<_E$. To do so, we define the following partial order on variables in the lower part of $\mathsf{KSh2c}$ proofs in HNFC.

DEFINITION 47
Let ϕ be a proof in HNFC. Define the partial order $<_\phi$ on the variables of occurring in $Lo(\phi)$ to be the minimal partial order such that $y <_\phi x$ if $K_1\{Q_1xK_2\{Q_2yB\}\}$ is a section of $Lo(\phi)$.

PROPOSITION 48
$<_\phi$ is well-defined for all $\mathsf{KSh2c}$ proofs in HNFC.

PROOF. Let ϕ be a proof of A in HNF, as in Definition 41. As $Lo(\phi)$ only contains $\mathsf{h}\!\downarrow, \mathsf{r1}\!\downarrow, \mathsf{r2}\!\downarrow$ and $\mathsf{qi}\!\uparrow$ rules and no α-substitution, if a variable v occurs in $Lo(\phi)$ then v occurs in $\forall xH^+_\phi(A)$. Notice also that none of $\mathsf{h}\!\downarrow, \mathsf{r1}\!\downarrow, \mathsf{r2}\!\downarrow$ or $\mathsf{qi}\!\uparrow$ can play the role of ρ in the following scheme:

$$\rho\frac{K\{Q_1v_1A_1\}\{Q_2v_2A_2\}}{K'\{Q_1v_1\{K''Q_2v_2B\}\}}.$$

Therefore, we observe that if $K_1\{Q_1xK_2\{Q_2yB\}\}$ is a section of $Lo(\phi)$, then $\forall xH^+_\phi(A)$ is of the form $L_1\{Q_1xL_2\{Q_2yC\}\}$, i.e. no dependencies can be introduced below $\forall xH^+_\phi(A)$. Thus, $x <_\phi y$ iff. $\forall xH^+_\phi(A)$ can be written $L_1\{Q_1xL_2\{Q_2yC\}\}$ for some $L_1\{\ \}, L_2\{\ \}$ and C and is therefore a well-defined partial order. □

LEMMA 49
Let ϕ be an $\mathsf{KSh2c}$ proof in HNFC and e' an $\forall$-edge in $\pi_1(\phi)$ that points to the $\exists$-edge e. If $Lab(e') = y$ and the $\exists$-node below e is $\exists xA$, then $x <_\phi y$.

PROOF. Since we have an $\exists$-node $\exists xA$ in $\pi_1(\phi)$ with an edge labelled t below it, there must be the following $\mathsf{h}\!\downarrow$ rule in ϕ:

$$K\left\{\mathsf{h}\!\downarrow\frac{\exists xA \vee [\tau \Rightarrow x]A}{\exists xA}\right\}$$

Since e points to e', y must occur freely in t. As ϕ is closed, y cannot be a free variable in $K\{\exists xA \vee [\tau \Rightarrow x]A\}$. Thus, $K\{\ \}$ must be of the form $K_1\{\forall yK_2\{\ \}\}$. Therefore, $x <_\phi y$. □

LEMMA 50
Let ϕ be an $\mathsf{KSh2c}$ proof in HNFC and e' an $\forall$-edge in $\pi_1(\phi)$ that points to the cut-edge e. If $Lab(e') = y$, E is the expansion tree below the cut edge e with $Sh(E) = A \wedge \bar{A}$) and Qx is some quantifier appearing in A (with $\bar{Q}x$ appearing in $\bar{A}$), then $x <_\phi y$.

PROOF. The cut-edge e in $\pi_1(\phi)$ corresponds to some cut $K\left\{\mathsf{qi}\!\uparrow\frac{A \wedge \bar{A}}{\mathsf{f}}\right\}$ in ϕ. Since e' points to e, we know that $y \in FV(A)$. But we also know that ϕ is a closed proof. Therefore, $K\{\ \} = K_1\{\forall yK_2\{\ \}\}$, and $x <_\phi y$. □

LEMMA 51
Let ϕ be an $\mathsf{KSh2c}$ proof in HNFC, e a $\forall$-edge of $\pi_1(\phi)$ labelled by x and e' an $\exists$-edge above an $\exists$-node $\exists yA$. If e is a descendant of e' then $x <_\phi y$.

PROOF. $Sh(\pi_1(\phi)) = K_1\{\exists y K_2 \forall x\{B\}\}$ (for some $K_1\{\ \}, K_2\{\ \}$, and B) is the conclusion of ϕ, so $x <_\phi y$. □

LEMMA 52
Let ϕ be an KSh2c proof in HNFC, $E_\phi = \pi_1(\phi)$ and e and e' be (not necessarily distinct) $\forall$-edges in E_ϕ s.t. $e <_{E_\phi} e'$, $Lab(e) = x$ and $Lab(e') = x'$. Then $x <_\phi x'$.

PROOF. As $e <_{E_\phi} e'$, there must be a chain

$$e_{q_0} <^-_{E_\phi} \cdots <^-_{E_\phi} e_{p_1} <^-_{E_\phi} e_{q_1} <^-_{E_\phi} \cdots <^-_{E_\phi} e_{p_m} <^-_{E_\phi} e_{q_m} <^-_{E_\phi} \cdots <^-_{E_\phi} e_{p_n}$$

where $e_{q_0} = e$ and $e_{p_n} = e'$, e_{q_i} points to e_{p_i}, and e_{q_i} is a descendant of $e_{p_{i+1}}$ in the expansion tree. If e_{q_i} points to e_{p_i}, then either e_{p_i} is an $\exists$-node or a cut node.

If e_{p_i} is an $\exists$-node, then by Lemma 49, we know that if $\exists x_{p_i}$ is the node above p_i and $Lab(e_{q_i}) = x_{q_i}$, then $x_{p_i} <_\phi x_{q_i}$. By Lemma 51, since $e_{q_{i-1}}$ is a descendant of e_{p_i} in the expansion tree, $x_{q_{i-1}} <_\phi x_{p_i}$, so we have $x_{q_{i-1}} <_\phi x_{q_i}$.

If e_{p_i} is a cut node, then we know that, since $e_{q_{i-1}}$ is a descendent of e_{p_i}, by Lemma 50, $x_{q_{i-1}} <_\phi x_{q_i}$.

Therefore, we have that $e_{q_0} <_\phi e_{q_{n-1}}$. Since $e_{q_{n-1}}$ must be a descendent of e_{p_n}, we have that $x = e_{q_0} <_\phi e_{p_n} = x'$. □

THEOREM 53
Let ϕ be a KSh2c proof of A in HNFC. Then we can construct an expansion proof with cut $E_\phi = \pi_1(\phi)$, with $Sh(E_\phi) = A$, and $Dp(E_\phi) = H_\phi(A)$.

PROOF. As described above, we only need to show that the dependency relation of E_ϕ is acyclic. Assume there were a cycle in $<_{E_\phi}$. Clearly, it could not be generated by just by travelling up the expansion tree. Thus, there is some $\forall$-edge e and an $\exists$-edge or cut edge e' such that e points to e' and $e <_{E_\phi} e' <_{E_\phi} e$. But then, if $Lab(e) = x$, by Lemma 52, $x <_\phi x$. But this contradicts Proposition 48. Therefore, $<_{E_\phi}$ is acyclic. □

5.2 *Expansion proofs to HNF*

For the translation from expansion proofs to KSh2c proofs in HNFC, we show that we can progressively build up a KSh2c by working through the 'minimal' nodes of an expansion proof. Unlike the previous translation, there is not necessarily a unique proof corresponding to each expansion proof, but a total order on universally quantified variables that respects $<_E$ is sufficient to give a unique proof up to equalities.

CONVENTION 54
We will not tend to omit 'with cut' in this section, as there are a few points where the distinction between cut-free and cut-full expansion proofs is important.

DEFINITION 55
A *weak* expansion tree is defined in the same way as in Definition 26 except that the first condition is weakened to allow any formula to be a leaf of the tree. A weak expansion tree with an acyclic dependency relation is correct regardless of whether its deep formula is a tautology.

A *weak* expansion tree with cut $E + \perp(E_1, \ldots, E_n)$ is just an expansion tree with cut where E and E_i are allowed to be weak expansion trees.

Definition 56

We define the *expansive deep formula* $Dp^+(E)$ for (weak) expansion trees, which is defined in the same way as the usual deep formula except that:

$$Dp^+(\exists xA +^{t_1} E_1 +^{t_2} \cdots +^{t_n} E_n) := \exists xA \vee Dp^+(E_1) \vee \cdots \vee Dp^+(E_n).$$

Definition 57

A *minimal edge* of a (weak) expansion tree (with cut) E is an edge that is minimal w.r.t. to $<_E$.

If all the edges below a node are minimal, we say that the node is a *minimal node*.

Lemma 58

If E is a weak expansion proof with cut with no minimal edges below existential nodes and no minimal universal and cut nodes, then it has a minimal $\star$ node.

Proof. Assume E is a weak expansion proof with no minimal edges below existential nodes and no minimal universal or cut nodes. Clearly, there must be must be at least one minimal edge e_0, and by the assumption it must be below a node $\star_0$. Let e'_0 be the other edge below $\star_0$. If e'_0 is minimal, we are done. If not, pick some minimal edge $e_1 < e'_0$, which again, with $e'_1 < e'_0$, must be below some $\star_1$. For each e'_i that is not minimal, we can find $e'_{i+1} < e'_i$. As E is finite, this sequence cannot continue indefinitely, so eventually we will find two minimal edges e_n and e'_n below $\star_n$. □

Lemma 59

Let $E = K_E\{\forall xA +^x A\}$, with $Dp^+(E) = K\{A\}$, be a correct weak expansion tree with a minimal $\forall$-edge labelled by x (which we will call e). Then there is a derivation $\begin{array}{c}\forall xK\{A\}\\ \Vert\{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow\}\\ K\{\forall xA\}\end{array}$.

Proof. We proceed by induction on the height of the node $\forall xA$ in E. If $\forall xA$ is the bottom node, then $K\{A\} = A$ and we are done. Let E be an expansion tree where $\forall x$ is not the bottom node. There are three possible cases to consider. In each case, $E_1 = K_{E_1}\{\forall xA +^x A\}$ is an expansion tree with $Dp^+(E_1) = K_1\{A\}$ and, by the inductive hypothesis, we have a derivation $\begin{array}{c}\forall xK_1\{A\}\\ \Vert\{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow\}\\ K_1\{\forall xA\}\end{array}$.

1. $E = (E_1 \star E_2)$, with $Dp^+(E) = K_1\{A\} \star Dp^+(E_2)$. As e is minimal, it cannot point to any edge in E_2. Therefore, $B := Dp^+(E_2)$ is free for x. Therefore, we can construct the derivations:

$$\mathsf{r1}\downarrow\frac{\forall x(K_1\{A\} \vee B)}{\begin{array}{c}\forall xK_1\{A\}\\ \Vert\{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow\} \vee B\\ K_1\{\forall xA\}\end{array}} \quad \text{and} \quad \mathsf{r2}\downarrow\frac{\forall x(K_1\{A\} \wedge B)}{\begin{array}{c}\forall xK_1\{A\}\\ \Vert\{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow\} \wedge B\\ K_1\{\forall xA\}\end{array}}$$

2. $E = \forall y(Sh(E_1)) +^y E_1$. As $Dp^+(E) = Dp^+(E_1)$, we are already done.
3. $E = \exists yK_0\{A_0\} +^{t_1} E_1 \cdots +^{t_n} E_n$, with $Dp^+(E_i) = B_i := [t_i \Rightarrow y](K_0\{A_0\})$ and in particular $B_1 = K_1\{A\}$. Thus, $Dp^+(E) = \exists yB_0 \vee K_1\{A\} \vee B_2 \vee \ldots \vee B_n$. Again, e cannot point to any edge in any

of the E_i', so we can construct:

$$\mathsf{r1}\!\downarrow \dfrac{\forall x(\exists y B_0 \vee K_1\{A\} \vee B_2 \vee \ldots \vee B_n)}{\mathsf{r1}\!\downarrow \dfrac{\forall x(\exists y B_0 \vee K_1\{A\})}{\left(\exists y B_0 \vee \begin{array}{c} \forall x K_1\{A\} \\ \Vert \{\mathsf{r1}\!\downarrow, \mathsf{r2}\!\downarrow\} \\ K_1\{\forall x A\} \end{array}\right)} \vee (B_2 \vee \ldots \vee B_n)}$$

□

LEMMA 60

Let $E = E_0 + \bot(E_1, \ldots, E_n)$, with $Dp^+(E) = Dp^+(E_0) \vee Dp^+(E_1) \vee \ldots \vee Dp^+(E_n)$, $Dp^+(E_i) = A_i$, and $A_k = K\{B\}$ for some particular $0 \leq k \leq n$, be a correct weak expansion tree with cut, s.t. the $\forall$-edge labelled by x (which we will call e) is minimal w.r.t. $<_E$.

Then there is a derivation:

$$\begin{array}{c} \forall x A_0 \vee A_1 \vee \ldots \vee K\{B\} \vee A_n \\ \Vert \{\mathsf{r1}\!\downarrow, \mathsf{r2}\!\downarrow\} \\ A_0 \vee A_1 \vee \ldots \vee K\{\forall x B\} \vee A_n \end{array}$$

PROOF. Since x is minimal w.r.t. $<_E$, it is certainly minimal w.r.t $<_{E_k}$. Therefore, by Lemma 59, we can construct the derivation $\begin{array}{c} \scriptstyle \forall x K\{B\} \\ \scriptstyle \Vert \{\mathsf{r1}\downarrow, \mathsf{r2}\downarrow\} \\ \scriptstyle K\{\forall x B\} \end{array}$. Therefore, we can also construct the derivation:

$$\begin{array}{c} \forall x(A_0 \vee A_1 \vee \ldots \vee A_k \vee \ldots \vee A_n) \\ \Vert \{\mathsf{r1}\!\downarrow\} \\ A_0 \vee A_1 \vee \ldots \vee \boxed{\begin{array}{c} \forall x K\{B\} \\ \Vert \{\mathsf{r1}\!\downarrow, \mathsf{r2}\!\downarrow\} \\ K\{\forall x B\} \end{array}} \vee \ldots \vee A_n \end{array}$$

□

DEFINITION 61

We define the map π_2^{Lo} : EPC $\to$ HNFC:

$$\pi_2^{Lo}(E) = \begin{array}{c} \forall \vec{x} Dp^+((E)) \\ \Vert \{\mathsf{h}\!\downarrow, \mathsf{r1}\!\downarrow, \mathsf{r2}\!\downarrow, qiU\} \\ Sh((E)) \end{array}$$

- If E is just a leaf A, $\pi_2^{Lo}(E) = A$.
- If $E = K_E\{B_1 \star_{E_1} C_1\} \ldots \{B_n \star_{E_n} C_n\}$, where E_i are all the $\star$-nodes s.t. the edges between $\star_{E_i}$ and B_i and between $\star_{E_i}$ and C_i are minimal, then we define $\pi_2^{Lo}(E) = E' = K_E\{B_1 \star_{F_1} C_1\} \ldots \{B_n \star_{F_n} C_n\}$, which is a correct weak expansion tree with cut. Pictorially:

$$E = K_E \left\{ \begin{array}{ccc} B_1 & & C_1 \\ & \diagdown \ \diagup & \\ & \star & \end{array} \right\} \cdots \left\{ \begin{array}{ccc} B_n & & C_n \\ & \diagdown \ \diagup & \\ & \star & \end{array} \right\}$$

$$E' = K_E\{B_1 \star C_1\} \ldots \{B_n \star C_n\}$$

- Assume E has no minimal $\star$ edges. If

$$E = K_E\{\exists x_1 A_1 +^{t_1^1} E_1^1 \cdots +^{t_1^{m_1}} E_1^{m_1}\}\ldots\{\exists x_n A_n +^{t_n^1} E_n^1 \cdots +^{t_n^{m_n}} E_n^{m_n}\}$$

with

$$Dp^+(E) = K\{\exists x A_1 \vee A_1^1 \vee \ldots \vee A_1^{m_1}\}\ldots\{\exists x A_n \vee A_n^1 \vee \ldots \vee A_n^{m_n}\}$$

and all edges $e_i^{j_i}$ minimal for $1 \leq i \leq n$ and $k_i < j_i \leq m_i$ (where $1 \leq k_i \leq m_i$), then

$$E' = K_E\{\exists x_1 A_1 +^{t_1^1} E_1^1 \cdots +^{t_1^{k_1}} E_1^{k_1}\}\ldots\{\exists x_n A_n +^{t_n^1} E_n^1 \cdots +^{t_n^{k_n}} E_n^{k_n}\}$$

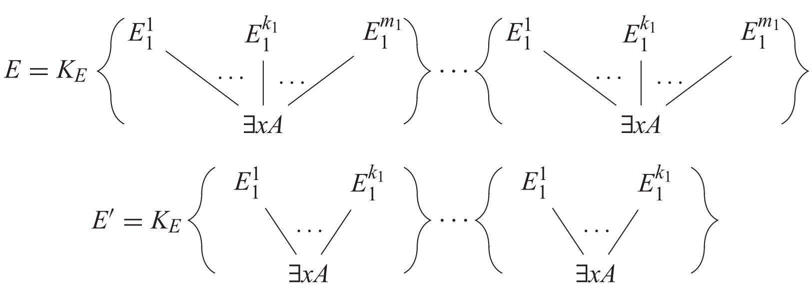
is a correct weak expansion tree with cut with

$$Dp^+(E) = K\{\exists x A_1 \vee A_1^1 \vee \ldots \vee A_1^{k_1}\}\ldots\{\exists x A_n \vee A_n^1 \vee \ldots \vee A_n^{k_n}\}$$

and we can define:

$$\pi_2^{Lo}(E) = \dfrac{K\left\{\begin{array}{c}\exists x A_1 \vee A_1^1 \vee \ldots \vee A_1^{m_1}\\ \| \{\mathsf{h}\downarrow\}\\ \exists x A_1 \vee A_1^1 \vee \ldots \vee A_1^{k_1}\end{array}\right\}\ldots\left\{\begin{array}{c}\exists x A_n \vee A_n^1 \vee \ldots \vee A_n^{m_n}\\ \| \{\mathsf{h}\downarrow\}\\ \exists x A_n \vee A_n^1 \vee \ldots \vee A_n^{k_n}\end{array}\right\}}{\pi_2^{Lo}(E')}{=}$$

Pictorially:



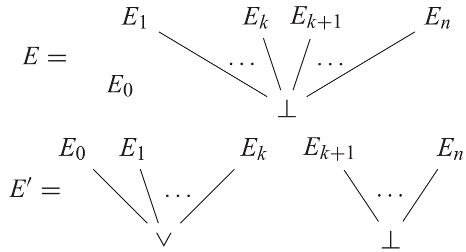
- Assume E has no minimal $\star$ nodes or $\exists$ edges. Let $E = E_0 + \bot(E_1, \ldots, E_k, \ldots, E_n)$ with the cut edges e_j for $1 \leq j \leq k$ all minimal. Then $E' = (E_0 \vee E_1 \vee \cdots \vee E_k) + \bot(E_{k+1}, \ldots, E_n)$ is a correct (weak) expansion tree with cut with

$$Sh(E') = (Sh(E_0) \vee Sh(E_1) \vee \cdots \vee Sh(E_k)) \text{ and } Dp^+(E') = Dp^+(E)$$

and we can define

$$\pi_2^{Lo}(E) = \dfrac{\begin{array}{c}\pi_2^{Lo}(E')\\ \hdashline A \vee \mathsf{qi}\uparrow\dfrac{A_1 \wedge \bar{A}_1}{\mathsf{f}} \vee \ldots \vee \mathsf{qi}\uparrow\dfrac{A_k \wedge \bar{A}_k}{\mathsf{f}}\end{array}}{A}{=}$$

Pictorially:



- Assume E is a weak expansion proof with cut with no minimal $\star$ or cut node, and no minimal $\exists$ edge. Then $E = K_E\{\forall xA +^x A\}$ for some minimal $\forall$ node, and by Lemma 60, $E' = K_E\{\forall xA\}$ is a correct weak expansion tree with cut and we can define:

$$\pi_2^{Lo}(E) = \begin{array}{c} \forall x Dp^+((E)) \\ \| \{\mathsf{r1}\downarrow, \mathsf{r2}\downarrow\} \\ Dp^+((E')) \\ = \overline{} \\ \pi_2^{Lo}(E') \end{array}$$

Pictorially,

$$E = K_E\left\{ \begin{array}{c} A \\ | \\ \forall x A \end{array} \right\} \qquad\qquad E' = K_E\{\forall x A\}$$

THEOREM 62
If E is an expansion proof with cut where $Sh(E) = A$, then we can construct an $\mathsf{KSh2c}$ proof ϕ of A in HNFC, where $H_\phi(A) = Dp(E)$.

PROOF. As $Dp(E)$ is a tautology, there is a proof $\begin{array}{c} \pi_2^{Up}(E) \Big\| \mathsf{KS} \\ \forall\vec{x}Dp(E) \end{array}$ and clearly there is a proof $\begin{array}{c} Dp(E) \\ \| \{\exists\mathsf{w}\downarrow\} \\ Dp^+(E) \end{array}$. Thus, assuming we have some strategy for picking minimal $\forall$-nodes, we can define π_2 from expansion proofs to $\mathsf{KSh2c}$ proofs in HNF as:

$$\pi_2(E) = \begin{array}{c} \pi_2^{Up}(E) \Big\| \mathsf{KS} \\ \forall\vec{x}Dp(E) \\ \| \{\exists\mathsf{w}\downarrow\} \\ \forall\vec{x}Dp^+(E) \\ \pi_2^{Lo}(E) \Big\| \{\mathsf{r1}\downarrow, \mathsf{r2}\downarrow, \mathsf{h}\downarrow, \mathsf{qi}\uparrow\} \\ Sh(E) \end{array}$$

□

OBSERVATION 63
For all expansion proofs with cut E we have $\pi_2^{Up}(E) = Up(\pi_2(E))$ and $\pi_2^{Lo}(E) = Lo(\pi_2(E))$.

REMARK 64
Although π_2 as defined here is a big improvement on the π_2 defined in [31], there is still a small element of choice involved. If one thinks game semantically, π_2 is equivalent to constructing a proof by ∃loise playing every possible move on her turn (it is fairly obvious that it doesn't make any significant difference in which order she makes these moves), followed by ∀belard choosing one possible move on his. Clearly which move ∀belard chooses affects the proof that will be constructed. Still, what we might call '∃loise canonicity' is an advance on what is possible in the sequent calculus, unless one adds some extra syntax, such as focussing [13].

We could make progress towards '∀belard canoncity' by replacing $\mathsf{r1}\downarrow$ and $\mathsf{r2}\downarrow$ with a general retract rule, such as in [10], but then we lose a certain amount of fine-grainedness in the proofs.

The translation for expansion proofs with closed cut is actually a lot more straightforward, since we can separate the cuts from $\mathsf{h}\downarrow, \mathsf{r1}\downarrow$ and $\mathsf{r2}\downarrow$.

COROLLARY 65
If E is an expansion proof with closed cuts s.t. $Sh(E) = A$, then we can construct a proof

$$\pi_3(E) = \begin{array}{c} \pi_3^{Up}(E) \Big\| \mathsf{KS} \\ \forall\vec{x}Dp(E) \\ \| \{\exists\mathsf{w}\downarrow\} \\ \forall\vec{x}Dp^+(E) \\ \pi_3^{Lo}(E) \Big\| \{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow,\mathsf{h}\downarrow\} \\ A \vee B \\ \| \{\mathsf{qi}\uparrow\} \\ A \end{array}$$

where $B = (\forall x A_1 \wedge \exists x A_1) \vee \ldots \vee (\forall x A_n \wedge \exists x A_n)$

PROOF. Instead of translating the expansion proof with cut, we replace the $\perp$ node with a series of $\vee$ nodes, to give an expansion proof E' with $Sh(E') = Sh(E) \vee (\forall x A_1 \wedge \exists x A_1) \vee \ldots \vee (\forall x A_n \wedge \exists x A_n)$ and $Dp(E') = Dp(E)$. Then, we just take

$$\pi_3(E) = \begin{array}{c} \pi_2(E') \Big\| \{\mathsf{h}\downarrow,\mathsf{r1}\downarrow,\mathsf{r2}\downarrow\} \\ \boxed{A \vee \mathsf{qi}\uparrow \dfrac{\forall x A_1 \wedge \exists x A_1}{\mathsf{f}} \vee \ldots \vee \mathsf{qi}\uparrow \dfrac{\forall x A_n \wedge \exists x A_n}{\mathsf{f}}} \end{array}$$

□

Of course, if the expansion proof is cut free, so is the deep inference proof.

COROLLARY 66
Let E be a cut-free expansion proof with $Sh(E) = A$. Then we can construct a proof ϕ_E in HNF of A.

PROOF. Clearly $\pi_2(E)$ is cut-free if E is. □

Having translations back and forth between expansion proofs and deep inference proofs gives us access to simple way to prove certain properties. For example, we can show eliminate switches from the lower part of HNF proofs.

PROPOSITION 67

The switch rule is admissible for the lower part of an HNF proof, i.e. if there is a proof in HNF,

$$\phi = \begin{array}{c} Up(\phi) \Vert \mathsf{KS} \\ H_\phi(A \wedge (B \vee C)) \\ \Vert \{\exists \mathsf{w}\downarrow\} \\ H^+_\phi(A \wedge (B \vee C)) \\ Lo(\phi) \Vert \{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow,\mathsf{h}\downarrow,\mathsf{qi}\uparrow\} \\ K\{A \wedge (B \vee C)\} \end{array}$$

we can construct:

$$\phi' = \begin{array}{c} Up(\phi') \Vert \mathsf{KS} \\ H_{\phi'}((A \wedge B) \vee C) \\ \Vert \{\exists \mathsf{w}\downarrow\} \\ H^+_{\phi'}((A \wedge B) \vee C) \\ Lo(\phi') \Vert \{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow,\mathsf{h}\downarrow,\mathsf{qi}\uparrow\} \\ K\{(A \wedge B) \vee C\} \end{array}$$

PROOF. Take $E_\phi = \pi_1(\phi)$. $E_\phi = K_E\{E_A \wedge (E_B \vee E_C)\}$, with $Sh(E_A) = A$, $Sh(E_B) = C$ and $Sh(E_C) = C$. Define $E' = K_E\{(E_A \wedge E_B) \vee E_C\}$. Clearly $Sh(E') = K\{[(A.B).C]\}$. We need to check if E' is correct. Clearly, any dependency cycle in E' could easily be transformed into a cycle in E. We have from ϕ a proof $\begin{array}{c}\Vert \mathsf{KS} \\ K'\{A' \wedge (B' \vee C')\}\end{array}$ where $A' = Dp(A)$, $B' = Dp(B)$ and $C' = Dp(C)$. Therefore, we have $Up(\phi') = \begin{array}{c}\Vert \mathsf{KS} \\ K'\left\{\mathsf{s}\dfrac{A' \wedge (B' \vee C')}{(A' \wedge B') \vee C'}\right\}\end{array}$, so E' is correct. Therefore, we can construct $\phi' = \pi_2(E')$. □

COROLLARY 68

$\mathsf{i}\uparrow$ is admissible for proofs in HNFC when in the lower part, i.e. if we have a proof

$$\phi = \begin{array}{c} \Vert \mathsf{KS} \\ H_\phi((A)) \\ \Vert \{\exists \mathsf{w}\downarrow\} \\ H^+_\phi((A)) \\ \Vert \{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow,\mathsf{h}\downarrow,\mathsf{qi}\uparrow,\mathsf{i}\uparrow\} \\ A \end{array}$$

we can construct a proof in HNFC

$$\phi' \begin{array}{c} \Vert \mathsf{KS} \\ H_{\phi'}((A)) \\ \Vert \{\exists \mathsf{w}\downarrow\} \\ H^+_{\phi'}((A)) \\ \Vert \{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow,\mathsf{h}\downarrow,\mathsf{qi}\uparrow\} \\ A \end{array}$$

PROOF. As seen previously $\mathsf{i}\uparrow$ is derivable for $\{\mathsf{ai}\uparrow, \mathsf{qi}\uparrow, \mathsf{s}\}$. We can simply push instances of $\mathsf{ai}\uparrow$ up through the lower part of the proof, and eliminate them in the upper part by propositional cut elimination.

By Proposition 67, we can eliminate any switches that are generated. □

6 Cut Elimination

We do not provide a direct cut elimination result for the deep inference proof systems. However, the translations to and from expansion proofs provide us with a number of 'off the shelf' procedures in the literature.

6.1 Cut elimination for expansion proofs

In the last decade a number of approaches to cut elimination for expansions have been developed. McKinley's Herbrand nets are proof nets for the sequent calculus, and so Herbrand net cut reductions adhere closely to those in the sequent calculus [29]. The cut reductions for Heijltjes's proof forests diverge from the sequent calculus, borrowing more from game semantical techniques [19]. However, a key ingredient for weak normalization is a different correctness condition to standard expansion trees, and thus it is not clear that some of the techniques made possible by this adjusted correctness condition—such as the *pruning* of proof forests—would translate naturally into either sequent calculus or deep inference. The unpublished work of Aschieri et al. gives a much more syntactic cut elimination procedure for Miller-style expansion proofs [6]. Unlike in McKinley and Heijltjes's papers, expansion trees are not limited to prenex formulae, although there is no *prima facie* reason why an extension to all first-order formulae would not be possible for these cut elimination procedures as well.

THEOREM 69
Let E be an expansion proof with cut. We can obtain a cut-free expansion proof E' with $Sh(E') = Sh(E)$.

PROOF. Using the techniques from [19], [29] or [6]. □

6.2 Cut elimination for SKSq

Making use of a cut elimination proof for expansion proofs, we can now prove an indirect cut elimination result for SKSq.

PROPOSITION 70
Let $\begin{array}{c}\scriptstyle \phi \,\Vert\, \mathsf{SKSq}\\ \scriptstyle A\end{array}$. Then we can construct $\begin{array}{c}\scriptstyle \phi' \,\Vert\, \mathsf{KSq}\cup\{\mathsf{i}\uparrow\}\\ \scriptstyle A\end{array}$, where every instance of $\mathsf{i}\uparrow$ is closed.

$$\begin{array}{c}\phi \,\Vert\, \mathsf{SKSq}\\ A\end{array} \longrightarrow \begin{array}{c}\phi' \,\Vert\, \mathsf{KSq}\cup\{\mathsf{i}\uparrow\}\\ A\end{array}$$

PROOF.

$$K\left\{\rho\uparrow\frac{A}{B}\right\} \longrightarrow {=}\frac{K\{A\}}{\mathsf{s}\dfrac{K\{A\}\wedge \mathsf{i}{\downarrow}\dfrac{\mathsf{t}}{\bar K\left\{\rho{\downarrow}\dfrac{\bar B}{\bar A}\right\}\vee K\{B\}}}{\mathsf{i}{\uparrow}\dfrac{K\{A\}\wedge\bar K\left\{\bar A\right\}}{\mathsf{f}}\vee K\{B\}}}$$

□

LEMMA 71

$$\begin{array}{c}\phi\,\Vert\,\mathsf{KSq}\cup\{\mathsf{i}\uparrow\}\\ A\end{array}\longrightarrow\begin{array}{c}\phi'\,\Vert\,\mathsf{KSq}\cup\{\mathsf{ai}\uparrow,\mathsf{qi}\uparrow\}\\ A\end{array}$$

with all $\mathsf{ai}\uparrow$, $\mathsf{qi}\uparrow$ closed.

PROOF. Straightforward. □

LEMMA 72
For any first-order formula context $K\{\ \}$ and any formula A, with no free variables bound by $K\{\ \}$, there are derivations

$$\begin{array}{c}K\{\mathsf{t}\}\wedge A\\ \Vert\{\mathsf{s},\mathsf{n}\uparrow,\mathsf{u}\uparrow\}\\ K\{A\}\end{array}\quad \textit{and} \quad \begin{array}{c}K\{A\}\\ \Vert\{\mathsf{s},\mathsf{n}\downarrow,\mathsf{u}\downarrow\}\\ K\{\mathsf{f}\}\vee A\end{array}$$

PROOF. We show that we can construct $\begin{array}{c}K\{\mathsf{t}\}\wedge A\\ \Vert\{\mathsf{s}\}\\ K\{A\}\end{array}$ by induction on the size of $K\{\ \}$. If $K\{\ \}=\{\ \}$ there is nothing to do.

The inductive steps are as follows:

$$\mathsf{s}\,\dfrac{\dfrac{K\{\mathsf{t}\}}{B\vee K'\{\mathsf{t}\}}\wedge A}{\dfrac{B\vee\boxed{\begin{array}{c}K'\{\mathsf{t}\}\wedge A\\ IH\,\Vert\{\mathsf{s},\mathsf{n}\uparrow,\mathsf{u}\uparrow\}\\ K'\{A\}\end{array}}}{K\{A\}}}\;,\qquad =\dfrac{\dfrac{K\{\mathsf{t}\}}{B\vee K'\{\mathsf{t}\}}\wedge A}{\dfrac{B\wedge\boxed{\begin{array}{c}K'\{\mathsf{t}\}\wedge A\\ IH\,\Vert\{\mathsf{s},\mathsf{n}\uparrow,\mathsf{u}\uparrow\}\\ K'\{A\}\end{array}}}{K\{A\}}}\;,\qquad =\dfrac{=\dfrac{K\{\mathsf{t}\}}{\forall xK'\{\mathsf{t}\}}\wedge A}{=\dfrac{\forall x\boxed{\begin{array}{c}\mathsf{n}\uparrow\dfrac{\forall xK'\{\mathsf{t}\}}{K'\{\mathsf{t}\}}\wedge A\\ IH\,\Vert\{\mathsf{s},\mathsf{n}\uparrow,\mathsf{u}\uparrow\}\\ K'\{A\}\end{array}}}{K'\{A\}}}\;,\ \text{and}$$

$$\mathsf{u}\uparrow\,\dfrac{=\dfrac{K'\{\mathsf{t}\}}{\exists xK'\{\mathsf{t}\}}\wedge=\dfrac{A}{\forall xA}}{=\dfrac{\exists x\boxed{\begin{array}{c}K'\{\mathsf{t}\}\wedge A\\ IH\,\Vert\{\mathsf{s},\mathsf{n}\uparrow,\mathsf{u}\uparrow\}\\ K'\{A\}\end{array}}}{K\{A\}}}$$

□

LEMMA 73
Let $\begin{array}{c}\phi\,\Vert\,\mathsf{KSq}\cup\{\mathsf{ai}\uparrow,\mathsf{qi}\uparrow\}\\ A\end{array}$ be a proof with all $\mathsf{ai}\uparrow$, $\mathsf{qi}\uparrow$ closed. Then we can construct a proof

$$\begin{array}{c}\phi'\,\Vert\,\mathsf{KSq}\cup\{\mathsf{ai}\uparrow\}\\ =\dfrac{A\vee\boxed{\mathsf{qi}\uparrow\dfrac{\exists xA_1\wedge\forall x\bar{A}_1}{\mathsf{f}}}\vee\ldots\vee\boxed{\mathsf{qi}\uparrow\dfrac{\exists xA_n\wedge\forall x\bar{A}_n}{\mathsf{f}}}}{A}\end{array}$$

where all the $\mathsf{ai}\uparrow$, $\mathsf{qi}\uparrow$ are still closed.

PROOF. We omit this proof, which can be found as the proofs of Lemmas 1.54 and 3.28 in [32]. Essentially, we can use Lemma 73 to push instances of $\mathsf{ai}\uparrow$ and $\mathsf{qi}\uparrow$ to the bottom of the proof. □

PROPOSITION 74
Every proof $\begin{smallmatrix}\phi \,\Vert\, \mathsf{SKSq}\\ A\end{smallmatrix}$ can be separated into a quantifier-cut-free top half, with parallel closed quantifier-cuts in the bottom half:

$$\begin{array}{c}\phi \,\Vert\, \mathsf{SKSq}\\ A\end{array} \longrightarrow \begin{array}{c}\Vert\, \mathsf{KSq}\cup\{\mathsf{ai}\uparrow\}\\ A'\\ \Vert\, \{\mathsf{qi}\uparrow\}\\ A\end{array}$$

PROOF. By the lemmas above. □

THEOREM 75
SKSq and KSq are weakly equivalent, i.e. if there is a proof $\begin{smallmatrix}\phi \,\Vert\, \mathsf{SKSq}\\ A\end{smallmatrix}$ then there is a proof $\begin{smallmatrix}\psi \,\Vert\, \mathsf{KSq}\\ A\end{smallmatrix}$.

PROOF. Let $\begin{smallmatrix}\phi \,\Vert\, \mathsf{SKSq}\\ A\end{smallmatrix}$. By Proposition 74 we can reduce all the up-rules to $\mathsf{ai}\uparrow$ and $\mathsf{qi}\uparrow$, pushing all instances of $\mathsf{qi}\uparrow$ to the bottom of the proof.

By Theorem 19, taking into account Remark 23, we can construct a Herbrand proof ϕ_2 of $A \wedge B$ where $B = (\forall x_1 B_1 \wedge \exists x_1 \bar{B}_1) \vee \ldots \vee (\forall x_n B_n \wedge \exists x_n \bar{B}_n)$.

By Theorem 12, we can eliminate $\mathsf{ai}\uparrow$ from the upper part of the Herbrand proof to form ϕ_3. By Proposition 44, we can construct a proof ϕ_4 of $A \wedge B$ in HNF.

By Theorem 53, we can construct an expansion proof with cut EC_{ϕ_4}. By Theorem 69, we can eliminate the cuts from EC_{ϕ_4} to give E_{ϕ_5}, and then translate back into a proof in HNF of A, ϕ_5 by Theorem 66. By Proposition 44, we can translate this back into a $\mathsf{KSh1}$ proof ϕ_6, and finally back into KSq with Proposition 24.

$$\begin{array}{ccccccc}
\begin{array}{c}\phi \,\Vert\, \mathsf{SKSq}\\ A\end{array} & \xrightarrow{\text{Prop } 74} & \begin{array}{c}\phi_1 \,\Vert\, \mathsf{KSq}\cup\{\mathsf{ai}\uparrow\}\\ A \wedge B\\ \Vert\, \{\mathsf{qi}\uparrow\}\\ A\end{array} & \xrightarrow{\text{Thm } 12} & \begin{array}{c}\phi_2 \,\Vert\, \mathsf{KSh1}\cup\{\mathsf{ai}\uparrow\}\\ A \wedge B\\ \Vert\, \{\mathsf{qi}\uparrow\}\\ A\end{array} & \xrightarrow{\text{Thm } 19} & \begin{array}{c}\phi_3 \,\Vert\, \mathsf{KSh1}\\ A \wedge B\\ \Vert\, \{\mathsf{qi}\uparrow\}\\ A\end{array}\\
 & \xrightarrow{\text{Prop } 44} & \begin{array}{c}\phi_4 \,\Vert\, \mathsf{KSh2}\\ A \wedge B\\ \Vert\, \{\mathsf{qi}\uparrow\}\\ A\end{array} & \xrightarrow{\text{Thm } 53} & \begin{array}{c}EC_{\phi_4} \,\Vert\, EPC\\ A\end{array} & \xrightarrow{\text{Thm } 69} & \begin{array}{c}E_{\phi_5} \,\Vert\, EP\\ A\end{array}\\
 & \xrightarrow{\text{Thm } 62} & \begin{array}{c}\phi_5 \,\Vert\, \mathsf{KSh2}\\ A\end{array} & \xrightarrow{\text{Prop } 44} & \begin{array}{c}\phi_6 \,\Vert\, \mathsf{KSh1}\\ A\end{array} & \xrightarrow{\text{Prop } 24} & \begin{array}{c}\psi \,\Vert\, \mathsf{KSq}\\ A\end{array}
\end{array}$$

□

7 Conclusion and Further Work

In this paper, we have brought together three strands of research together—the study of Herbrand proofs, the study of expansion proofs, and deep inference proof theory—and shown how the notion of decomposed proofs from deep inference proof theory can be used to unify the two approaches to Herbrand theorem. The theory of decomposition for first-order classical logic and other logics is still being developed, led by Aler Tubella [4, 5], and any general advances would be likely to have a bearing on this line of research.

However, we are not wedded to the deep inference methodology, and, in fact, there are other attractive ways of representing decomposed proofs. In particular the use of *combinatorial proofs* to represent decomposed proofs is an interesting development [24, 34]. Recently, Hughes has introduced first-order combinatorial proofs, and work is being carried out by the author and others to fit these graph-theoretical proof objects into the picture [25]. Combinatorial proofs provide natural, 'syntax-free' equivalence classes for decomposed proofs in deep inference, and already the author has investigated how variants of propositional combinatorial proofs can be used to classify sets of structural deep-inference inference rules. Already, investigations have begun in extending this work to first-order classical logic, a line of research motivated by many of the same logical and philosophical concerns that animated Herbrand's own investigations almost a century ago.

Acknowledgements

The author would like to thank the anonymous reviewer for their insightful and helpful comments and to acknowledge support from the Engineering and Physical Sciences Research Council Project EP/K018868/1 Efficient and Natural Proof Systems and the Agence Nationale de la Recherche project FISP ANR-15-CE25-0014-01.

References

[1] J. P. Aguilera and M. Baaz. Unsound Inference Make Proofs Shorter. *The Journal of Symbolic Logic*, **84**, 102–122, 2019.

[2] A. Alcolei, P. Clairambault, M. Hyland and G. Winskel. The True Concurrency of Herbrand's Theorem. In *27th EACSL Annual Conference on Computer Science Logic (CSL 2018)*, volume 119 of Leibniz International Proceedings in Informatics (LIPIcs), D. Ghica and A. Jung, eds, pp. 5:1–5:22. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2018.

[3] A. A. Tubella. *A Study of Normalisation through Subatomic Logic*. *PhD Thesis*, University of Bath, 2017.

[4] A. A. Tubella and A. Guglielmi. Subatomic Proof Systems: Splittable Systems. *ACM Transactions on Computational Logic*, **19**, 1–33, 2018.

[5] A. A. Tubella, A. Guglielmi and B. Ralph. Removing Cycles from Proofs. In *26th EACSL Annual Conference on Computer Science Logic (CSL 2017)*. Volume 82 of Leibniz International Proceedings in Informatics (LIPIcs), V. Goranko and M. Dam, eds, pp. 9:1–9:17. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017.

[6] F. Aschieri, S. Hetzl and D. Weller. Expansion Trees with Cut. *Mathematical Structures in Computer Science*, **29**, 1009–1029, 2018.

[7] M. Baaz and A. Leitsch. On skolemization and proof complexity. *Fundamenta Informaticae*, **20**, 353–379, 1994.

[8] K. Brünnler. *Deep Inference and Symmetry in Classical Proofs*. Logos, 2003.

[9] K. Brünnler. Two restrictions on contraction. *Logic Journal of IGPL*, **11**, 525–529, 2003.

[10] K. Brünnler. Cut elimination inside a deep inference system for classical predicate logic. *Studia Logica*, **82**, 51–71, 2006.

[11] K. Brünnler and A. Tiu. A local system for classical logic. In *Logic for Programming, Artificial Intelligence, and Reasoning*, vol. 2250, pp. 347–361. Springer, 2001.

[12] S. R. Buss. On Herbrand's theorem. In *Logic and Computational Complexity*, pp. 195–209. Springer, 1995.

[13] K. Chaudhuri, S. Hetzl and D. Miller. A multi-focused proof system isomorphic to expansion proofs. *Journal of Logic and Computation*, **26**, 577–603, 2016.

[14] A. Degtyarev and A. Voronkov. Kanger's choices in automated reasoning. In *Collected Papers of Stig Kanger with Essays on His Life and Work*, pp. 53–67. Springer, 2001.
[15] B. Dreben, P. Andrews and S. Aanderaa. False lemmas in Herbrand. *Bulletin of the American Mathematical Society*, **69**, 699–706, 1963.
[16] A. Guglielmi. A system of interaction and structure. *ACM Transactions on Computational Logic*, **8**, 2007.
[17] A. Guglielmi, T. Gundersen and M. Parigot. A Proof Calculus Which Reduces Syntactic Bureaucracy. *Proceedings of the 21st International Conference on Rewriting Techniques and Applications (RTA'10)*. **6**, 135–150, 2010.
[18] T. E. Gundersen. A General View of Normalisation through Atomic Flows. *PhD Thesis*, University of Bath, Bath, 2009.
[19] W. Heijltjes. Classical proof forestry. *Annals of Pure and Applied Logic*, **161**, 1346–1366, August 2010.
[20] J. Herbrand. Recherches Sur La Théorie de La Démonstration. *PhD Thesis*, University of Paris, 1930.
[21] J. Herbrand. *Logical Writings*. Springer, 1971.
[22] D. Hilbert. On the Infinite. Text of an address delivered in Münster on 4th June 1925 at a meeting of the Westphalian Mathematical Society. *Jean van Heijenoort*, 1878–1931, 1967.
[23] D. Hilbert and P. Bernays. *Grundlagen Der Mathematik II*. Springer, 1974.
[24] D. J. D. Hughes. *Towards Hilbert's 24th Problem: Combinatorial Proof Invariants (Preliminary version)*. Electronic Notes in Theoretical Computer Science, **165**, 37–63, 2006.
[25] D. J. D. Hughes. First-order proofs without syntax. arXiv:1906.11236 [math], June 2019.
[26] S. Kanger. Provability in logic. In *Stockholm studies in Philosophy*, vol. 1. Alynquist & Wiksell, 1957.
[27] R. Kuzmin. Sur une nouvelle classe de nombres transcendants. *Bulletin de l'Académie des Sciences de l'URSS. VII série*, 585–597, 1930.
[28] R. McKinley. A sequent calculus demonstration of Herbrand's theorem. Preprint, arXiv:1007.3414, 2010.
[29] R. McKinley. Proof nets for Herbrand's Theorem. *ACM Transactions on Computational Logic (TOCL)*, **14**, 1–31, 2013.
[30] D. A. Miller. A compact representation of proofs. *Studia Logica*, **46**, 347–370, 1987.
[31] B. Ralph. A Natural Proof System for Herbrand's Theorem. In *International Symposium on Logical Foundations of Computer Science*, pp. 289–308. Springer, 2018.
[32] B. Ralph. *Modular Normalisation of Classical Proofs*. PhD Thesis, University of Bath, 2019. Accepted: 2019.
[33] R. Statman. Lower bounds on Herbrand's theorem. *Proceedings of the American Mathematical Society*, 104–107, 1979.
[34] L. Straßburger. Combinatorial flows and their normalisation. In *LIPIcs-Leibniz International Proceedings in Informatics*, vol. 84. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017.
[35] L. Straßburger. The problem of proof identity, and why computer scientists should care about Hilbert's 24th problem. *Philosophical Transactions of the Royal Society A*, **377**, 20180038, 2019.
[36] R. Thiele. Hilbert's twenty-fourth problem. *American Mathematical Monthly*, **110**, 1–24, 2003.
[37] J. Webb. *On Herbrand's Theorem*. Master's Thesis, University of Oxford, 2014.

Received 25 March 2020