NOTES FOR SEQUENT CALCULUS PRESENTATION

MILES KRETSCHMER

Contents

1. What is a sequent?

The sequent calculus is a proof system for first order logic, $L_{\omega,\omega}$. We assume that the language contains infinitely many constant symbols, which can be treated as implicitly universally quantified free variables. A derivation in the sequent calculus consists of **sequents**, which are expressions of the form

$$\varphi_1, \ldots, \varphi_m \vdash \psi_1, \ldots, \psi_n$$

where the $\varphi_i, \psi_j \in L_{\omega,\omega}$. One can read this expression as "if all the φ_i are true, then at least one of the ψ_j is true." A classically equivalent reading which may be helpful for understanding certain rules in the calculus is "either one of the φ_i is false, or one of the ψ_j is true." This reading underscores the symmetry of what a sequent is expressing. The sequent is giving us a list of positive and negative possibilities.

A proof of a sentence φ will be interpreted as a derivation of the sequent

$$\vdash \varphi$$

The empty sequent

$$\vdash$$

is the sequent calculus's version of a formal contradiction. There are no possibilities!

For a bit of motivation, the sequent calculus differs from other proof systems in that each step consists of a conditional assertion, with the conditions explicitly included as part of the statement. This allows us to replace elimination rules with rules that add to the left of the sequent. With the exception of the (eliminable) cut rule (and perhaps equality rules), all rules in the sequent calculus add, rather than remove, syntactic

Date: April 2024.

material to the sequent. This makes (cut-free) derivations in the sequent calculus more analyzable than other kinds of derivations.

2. Derivations in the Sequent Calculus

A derivation in the calculus is a tree of sequents. At the bottom, we have the sequent we are deriving. Above each sequent is one or two sequents, such that a rule of inference yields the sequent below from the sequents above.

2.1. Axioms. At the leaves of the tree are logical axioms, sequents of the form

$$\varphi \vdash \varphi$$

(We may also wish to include $\vdash \top$ and $\bot \vdash$ as logical axioms.)

2.2. Structural Rules. The structural rules are those rules which operate on sequents without regard to the internal structure of their constituent formulas. All rules in the sequent calculus except the cut rule will come in pairs, one operating on the left and one on the right.

The weakening rules allow us to add formulas to a sequent. Adding more possibilities always makes things more true.

$$\frac{\Gamma \vdash \Delta}{\varphi, \Gamma \vdash \Delta}\ \mathrm{WL} \qquad \frac{\Gamma \vdash \Delta}{\Gamma \vdash \Delta, \varphi}\ \mathrm{WR}$$

The contraction rule collapses multiple instances of the same formula.

$$\frac{\varphi, \varphi, \Gamma \vdash \Delta}{\varphi, \Gamma \vdash \Delta}\ \mathrm{CL} \qquad \frac{\Gamma \vdash \Delta, \varphi, \varphi}{\Gamma \vdash \Delta, \varphi}\ \mathrm{CR}$$

The interchange rule permutes formulas on one side of the sequent.

$$\frac{\Gamma, \varphi, \psi, \Theta \vdash \Delta}{\Gamma, \psi, \varphi, \Theta \vdash \Delta}\ \mathrm{IL} \qquad \frac{\Gamma \vdash \Delta, \varphi, \psi, \Lambda}{\Gamma \vdash \Delta, \psi, \varphi, \Lambda}\ \mathrm{IR}$$

The cut rule is considered a structural rule, but we will address it later due to its special status.

I will generally pass over these rules without comment, treating the sides of the sequent as sets of formulas, but I will mention in passing that although these rules seem to be entirely innocent, some interesting things come from restricting the contraction rule (linear logic), the weakening rule (minimal logic), and even the interchange rule (Lambek calculus).

2.3. Connective Rules. I will now address the logical rules, which modify the structure of the constituent formulas. Here we record the rules for connectives.

$$\frac{\varphi, \psi, \Gamma \vdash \Delta}{\varphi \wedge \psi, \Gamma \vdash \Delta}\ \wedge\mathrm{L} \qquad \frac{\Gamma \vdash \Delta, \varphi \qquad \Gamma \vdash \Delta, \psi}{\Gamma \vdash \Delta, \varphi \wedge \psi}\ \wedge\mathrm{R}$$

$$\frac{\varphi, \Gamma \vdash \Delta \qquad \psi, \Gamma, \vdash \Delta}{\varphi \vee \psi, \Gamma \vdash \Delta}\ \vee\mathrm{L} \qquad \frac{\Gamma \vdash \Delta, \varphi, \psi}{\Gamma \vdash \Delta, \varphi \vee \psi}\ \vee\mathrm{R}$$

Small note: there is another way to formulate $\wedge$L and $\vee$R, where one of the formulas is not required to be present already. This makes no difference in the presence of the weakening rules, but in intuitionistic logic and minimal logic one uses this other formulation. I will use it later.

$$\frac{\Gamma \vdash \Delta, \varphi \qquad \psi, \Theta \vdash \Lambda}{\varphi \to \psi, \Gamma, \Theta \vdash \Delta, \Lambda}\ \to\mathrm{L} \qquad \frac{\varphi, \Gamma \vdash \Delta, \psi}{\Gamma \vdash \Delta, \varphi \to \psi}\ \to\mathrm{R}$$

$$\frac{\Gamma \vdash \Delta, \varphi}{\neg\varphi, \Gamma \vdash \Delta}\ \neg\mathrm{L} \qquad \frac{\varphi, \Gamma \vdash \Delta}{\Gamma \vdash \Delta, \neg\varphi}\ \neg\mathrm{R}$$

Before we proceed with the quantifier rules, here is an example derivation using the connectives.

$$\dfrac{\dfrac{\dfrac{p \vdash p}{\vdash p, \neg p}\,\neg\mathrm{R} \qquad q \vdash q}{p \to q \vdash \neg p, q}\,{\to}\mathrm{L}}{p \to q \vdash \neg p \vee q}\,\vee\mathrm{R}$$

2.4. **Quantifier Rules.** The quantifier rules require some special attention to the occurrence of variables.

$$\dfrac{\varphi(t), \Gamma \vdash \Delta}{\forall x \varphi(x), \Gamma \vdash \Delta}\,\forall\mathrm{L} \qquad \dfrac{\Gamma \vdash \Delta, \varphi(c)}{\Gamma \vdash \Delta, \forall x \varphi(x)}\,\forall\mathrm{R}$$

The $\forall$R requires c to be a constant unused in Γ and Δ.

$$\dfrac{\varphi(c), \Gamma \vdash \Delta}{\exists x \varphi(x), \Gamma \vdash \Delta}\,\exists\mathrm{L} \qquad \dfrac{\Gamma \vdash \Delta, \varphi(t)}{\Gamma \vdash \Delta, \exists x \varphi(x)}\,\exists\mathrm{R}$$

The $\exists$L rules requires c to be an unused constant as above.

2.5. **The Cut Rule.** The cut rule is as follows

$$\dfrac{\Gamma \vdash \Delta, \varphi \qquad \varphi, \Theta \vdash \Lambda}{\Gamma, \Theta \vdash \Delta, \Lambda}\,\mathrm{CUT}$$

The formula φ is called the cut formula. Notice that this rule is the only one in which some syntactic material which appears above the line is truly destroyed. This property is responsible for the special status of the cut rule.

Here is an example of a derivation using the cut rule and the quantifier rules:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\neg P(c) \vdash \neg P(c)}{\neg P(c) \vdash \exists x \neg P(x)}\,\exists\mathrm{R}}{\vdash \exists x \neg P(x), \neg\neg P(c)}\,\neg\mathrm{R} \qquad \dfrac{\dfrac{P(c) \vdash P(c)}{\vdash P(c), \neg P(c)}\,\neg\mathrm{R}}{\neg\neg P(c) \vdash P(c)}\,\neg\mathrm{L}}{\vdash \exists x \neg P(x), P(c)}\,\mathrm{CUT}}{\vdash \exists x \neg P(x), \forall x P(x)}\,\forall\mathrm{R}}{\neg \forall x P(x) \vdash \exists x \neg P(x)}\,\neg\mathrm{L}$$

2.6. **Equality Rules and Non-logical Axioms.** We can add axioms to the sequent calculus in addition to the logical axioms $\varphi \vdash \varphi$, by allowing the associated sequent to appear at the top of a branch in a derivation. For the sake of things, I will list the equality axioms in sequent style.

$$\vdash t = t$$

$$t = s \vdash s = t$$
$$t = s, s = r \vdash t = r$$
$$t_1 = s_1, \ldots, t_n = s_n, R(\bar{t}) \vdash R(\bar{s})$$
$$t_1 = s_1, \ldots, t_n = s_n, \vdash f(\bar{t}) = f(\bar{s})$$

A problem with this approach is that whenever we have nonlogical axioms, the proof of the cut elimination theorem does not automatically apply. In particular, if $\vdash p$ and $p \vdash$ are axioms, I have the proof

$$\dfrac{\vdash p \qquad p \vdash}{\vdash}\,\mathrm{CUT}$$

There is no cut free proof of a contradiction, even using these axioms. What cut elimination gets us in general is a proof whose only cuts are applied to formulas which can be traced back to an axiom. This is called the free cut elimination theorem.

Another approach is to replace axioms with non-logical rules of inference. See below for the equality axioms:

$$\dfrac{t = t, \Gamma \vdash \Delta}{\Gamma \vdash \Delta,}\,\mathrm{refl} \qquad \dfrac{\bar{t} = \bar{s}, R(\bar{t}), R(\bar{s}), \Gamma \vdash \Delta}{\bar{t} = \bar{s}, R(\bar{t}), \Gamma \vdash \Delta}\,\mathrm{sub} \qquad \dfrac{\bar{t} = \bar{s}, f(\bar{t}) = f(\bar{s}), \Gamma \vdash \Delta}{\bar{t} = \bar{s}, \Gamma \vdash \Delta}\,\mathrm{sub}$$

Doing this, we can recover cut elimination provided these rules have a constrained form, like the above. The trade-off is that cut elimination tells us less, because cut free proofs now potentially contain these inferences. The equality rules are still fairly analyzable (especially when we have no function symbols), so they will be ok. I will not consider them in the proof of the cut elimination theorem, however.

3. Duality and Subsystems

There is a lot of symmetry in how the sequent calculus works. I will now digress to describe some aspects of this. Each logical symbol has a dual symbol: $\neg$ is its own dual, $\wedge$ corresponds to $\vee$, $\forall$ to $\exists$, and $\to$ can correspond to $\not\leftarrow$ if we want to include that (also $\top$ corresponds to $\bot$). By replacing all instances of a logical symbol with its dual symbol, we can translate a formula φ to its dual formula $\varphi^\star$, likewise a list of formulas Γ to a list $\Gamma^\star$. We can then define for a sequent

$$\sigma : \Gamma \vdash \Delta$$

the dual sequent

$$\sigma^\star : \Delta^\star \vdash \Gamma^\star$$

Theorem. *If a sequent σ is derivable, its dual $\sigma^\star$ is also derivable.*

Proof. Given a derivation π of σ, we can dualize everything in π to get a tree of sequents $\pi^\star$ ending in $\sigma^\star$. It remains to check that $\pi^\star$ is actually a valid derivation, or in other words, that every step of $\pi^\star$ is a valid rule of inference. We can do this by observing that every rule has a dual rule, whose premises and conclusions are respectively the duals of the original rule. (Exercise: Formulate rules for $\not\leftarrow$ to make this true when that is included, and convince yourself that they make sense.) The cut rule is self dual. $\square$

3.1. **Intuitionistic Logic.** The above duality rests on the duality pertaining to negation in classical logic. Breaking this symmetry, we obtain an intuitionistic subsystem of the sequent calculus.

We say a sequent is *intuitionistic* if it contains at most one formula on the right side. A derivation is intutionistic if it consists only of intuitionistic sequents. To make this work properly, we need to modify the $\vee$R rule so that it only needs one of the disjuncts to be present. This is no big deal.

$$\frac{\Gamma \vdash \varphi}{\Gamma \vdash \varphi \vee \psi}\ \vee\mathrm{Rr} \qquad \frac{\Gamma \vdash \varphi}{\Gamma \vdash \psi \vee \varphi}\ \vee\mathrm{Rl}$$

As a nonexample, we have the following classical derivations:

$$\dfrac{\dfrac{p \vdash p}{\vdash p, \neg p}\ \neg\mathrm{R}}{\vdash p \vee \neg p}\ \vee\mathrm{R} \qquad \dfrac{\dfrac{\dfrac{p \vdash p}{\vdash p, \neg p}\ \neg\mathrm{R}}{\neg\neg p \vdash p}\ \neg\mathrm{L}}{\vdash \neg\neg p \to p}\ \to\mathrm{R}$$

As we would hope, these are not intuitionistic.

It is certainly not obvious that this strange condition on sequents is what captures intuitionistic provability. Formally this can be justified by a translation to another intuitionistic proof system. I also invite you to contemplate a BHK interpretation of the intuitionistic calculus, where we construct a function taking proofs of the formulas on the left to a proof of the formula on the right. In this case, the intuitionistic restriction makes the negation and implication rules interpretable in this way. I will also be able to explain this somewhat after we have the cut elimination theorem.

3.2. **Dual Intuitionistic Logic.** A feature of this definition is that it can be dualized. We say a sequent is *dual intuitionistic* if it contains at most one formula on the left, and a derivation is dual intuitionistic if it contains only dual intuitionistic sequents. (We also need to modify $\wedge$L) An example of a derivation that is not dual intuitionistic is

$$\dfrac{\dfrac{\dfrac{\dfrac{p \vdash p}{\neg p, p \vdash}\ \neg\mathrm{L}}{\neg p \wedge p \vdash}\ \wedge\mathrm{L}}{\neg p \wedge p \vdash q}\ \mathrm{WR}}{\vdash \neg p \wedge p \to q}\ \to\mathrm{R}$$

The sequent at the end is not derivable in dual intuitionistic logic, this is a paraconsistent logic.

4. Cut Free Proofs and The Subformula Property

We now return to the main program of the talk. We say that a proof is cut free if it does not use the cut rule. As we mentioned previously, the cut rule is the only rule in the sequent calculus that destroys syntactic material (ignoring equality rules). We can now make this precise.

A proof of $\Gamma \vdash \Delta$ is said to have the *subformula property* if for any sequent $\Theta \vdash \Lambda$ which appears in the proof, and any φ appearing in Θ, Λ, φ is a subformula of some formula appearing in Γ, Δ.

Proposition. *If π is a cut free proof, π has the subformula property.*

Proof. Induction on the structure of π. If π is an axiom, it has the subformula property. Otherwise, π ends in a non-cut inference of σ from one or two premises σ_i, each of which is the end of a shorter cut-free proof π_i. We call these the left and right branches if there are two of them. Appealing to induction each π_i has the subformula property, so any formula appearing in π_i is a subformula of one appearing in σ_i. Now to check that π has the subformula property we examine each possible inference that could have lead to σ and verify that every formula in σ_i persists as a subformula of a formula in σ. □

To give a preview of the usefulness of cut-free proofs, I record the following.

Proposition. *There is no cut free proof of a contradiction, the empty sequent $\vdash$.*

Proof. There are no subformulas that could appear in previous steps of such a proof. □

5. Cut Elimination

Our means of obtaining cut free proofs is the following.

Theorem (Cut Elimination)**.** *Let σ be a derivable sequent. There is a cut free proof π of σ. In fact, such a proof can be obtained by an algorithmic process from a cut-full proof of σ, and is intuitionistic if the original proof was intuitionistic.*

I will now give a sketch of the proof. For technical reasons, we will want to modify the cut rule to make it easier to analyze. Our new cut rule has the form

$$\frac{\Gamma \vdash \Delta \qquad \Theta \vdash \Lambda}{\Gamma, [\Theta \setminus \varphi] \vdash [\Delta \setminus \varphi], \Lambda} \text{ CUT}$$

where φ occurs in Δ and Θ (anywhere), and $[\Delta \setminus \varphi]$, $[\Theta \setminus \varphi]$, are obtained by deleting all instances of φ. In the presence of the structural rules, this is equivalent to the ordinary cut rule. Essentially, we are stipulating that the rule act on all instances of the cut formula simultaneously, and ignoring structural issues.

The proof has a somewhat complex inductive structure, which I will try to separate out as clearly as possible. There is a local problem, reducing the complexity of a single cut, and a global one, eliminating all cuts from the proof. I will begin by laying out the inductive structure that reduces the global problem to the local one.

5.1. The Global Problem.

We define the *rank* of a cut to be the rank of the cut formula. We define the total rank of a proof to be the least integer greater than the rank of all cuts appearing in the proof. Thus a proof of total rank 0 is cut-free. The global structure of the argument is an induction on total rank. We are thus proving the following statement by induction on n.

Proposition (cut elimination: n)**.** *If π is a proof of a sequent σ of total rank at most n, there is a cut-free proof π' of σ.*

For $n = 0$, this is trivial, a proof of total rank 0 is cut-free. Our task now is to prove cut elimination: $n+1$ assuming cut elimination: n. We break up the proof into a series of sublemmas which are proved under this assumption. The place where the local problem will be solved is in the first one:

Lemma (Reducing rank above a single cut: n)**.** *Suppose π is a proof of σ ending in a single rank at most n cut and containing no other instances of cut. Then there is a proof π' of σ of total rank n. In other words, π' contains only cuts of rank below n.*

We will postpone the proof of this to finish the global structure of the proof. This lemma, under our assumptions, immediately implies the following.

Lemma (cut elimination above a single cut: n)**.** *Suppose π is a proof of σ ending in a single rank at most n cut and containing no other cuts. Then there is a cut free proof π' of σ.*

Proof. First apply the lemma on reducing rank above a single cut: n to the proof to get a rank n proof. Now appealing to induction and using cut elimination : n, we have a cut-free proof. □

Proof of cut elimination: $n+1$. We argue by induction on the total number of cuts in the proof. If it is 0 we are done. If it is nonzero, find a cut with no cuts above it, and apply cut elimination above a single cut: n to replace this subproof with a cut free proof. This reduces the total number of cuts by 1. Appealing to induction we are done. □

5.2. **The Local Problem.** At this point the global inductive structure is complete, and we have reduced the global problem to its local form, reducing rank above a single cut: n.

Proof sketch: Reducing rank above a single cut: n. This is also an induction, but on a measure of depth of the cut below the introduction of the cut formula. We are looking at a proof π of the following form

$$\pi = \begin{cases} \vdots\,\pi_1 \quad \vdots\,\pi_2 \\ \dfrac{\sigma_1 \qquad \sigma_2}{\sigma}\ \text{CUT} \end{cases}$$

We define the left depth of the proof to be the the height above σ_1 at which the cut formula first appears in π_1, and the right depth likewise. In other words, this is how much time has elapsed between introducing the cut formula and applying cut. The total depth is the sum of the left and right depths. The remainder of the proof has two parts. First, there is the case that the total depth is zero. Here the cut formula is introduced immediately above the cut. We will indicate how to either (1) eliminate the cut entirely or (2) replace it with a cut on a subformula of the cut formula. Both of these solve the local problem. Then, there is the case of positive total depth. We will indicate how to (3) push the cut up in the tree, reducing the total depth.

5.2.1. *Simplifying a cut.* If the cut formula was introduced as an axiom or by weakening, we don't need to know anything about the other branch.

Axiom case: The cut formula is introduced as a logical axiom. We will completely eliminate it. For concreteness suppose our proof looks like this

$$\dfrac{\varphi \vdash \varphi \qquad \begin{matrix}\vdots\,\pi_2\\ \Gamma \vdash \Delta\end{matrix}}{\varphi, [\Gamma \setminus \varphi] \vdash \Delta}\ \text{CUT}$$

with the axiom on the left. $\varphi, [\Gamma \setminus \varphi]$ is equivalent to Γ up to the use of structural rules, so we can remove the CUT, replacing π with π_2. (In the case that the axiom is $\vdash \top$, one needs to prove by a straightforward induction that if there is a cut free proof of $\Gamma \vdash \Delta$, there is a cut free proof of $[\Gamma \setminus \top] \vdash \Delta$) The case that the axiom on the right is similar (dual).

Weakening case: The cut formula is introduced by weakening. If it is on the left we can make the following transformation:

$$\dfrac{\dfrac{\begin{matrix}\vdots\,\pi_1\\ \Gamma \vdash \Delta\end{matrix}}{\Gamma \vdash \Delta, \varphi}\ \text{WR} \qquad \begin{matrix}\vdots\,\pi_2\\ \Theta \vdash \Lambda\end{matrix}}{\Gamma, [\Theta \setminus \varphi] \vdash \Delta, \Lambda}\ \text{CUT} \quad \rightsquigarrow \quad \dfrac{\begin{matrix}\vdots\,\pi_1\\ \Gamma \vdash \Delta\end{matrix}}{\Gamma, [\Theta \setminus \varphi] \vdash \Delta, \Lambda}\ \text{WR}$$

as before, if it is on the right the situation is similar.

Example connective case:

$$\dfrac{\dfrac{\overset{\vdots\;\pi_1}{\Gamma\vdash\Delta,\varphi} \qquad \overset{\vdots\;\pi_2}{\Gamma\vdash\Delta,\psi}}{\Gamma\vdash\Delta,\varphi\wedge\psi}\wedge\mathrm{R} \qquad \dfrac{\overset{\vdots\;\pi_3}{\varphi,\psi,\Theta\vdash\Lambda}}{\varphi\wedge\psi,\Theta\vdash\Lambda}\wedge\mathrm{L}}{\Gamma,\Theta\vdash\Delta,\Lambda}\,\mathrm{CUT} \;\rightsquigarrow\; \dfrac{\overset{\vdots\;\pi_2}{\Gamma\vdash\Delta,\psi} \qquad \dfrac{\overset{\vdots\;\pi_1}{\Gamma\vdash\Delta,\varphi}\qquad \overset{\vdots\;\pi_3}{\varphi,\psi,\Theta\vdash\Lambda}}{\psi,\Gamma,\Theta\vdash\Delta,\Lambda}\,\mathrm{CUT}}{\Gamma,\Theta\vdash\Delta,\Lambda}\,\mathrm{CUT}$$

We have replaced one cut with two cuts on simpler formulas, solving the local problem.

Example quantifier case:

$$\dfrac{\dfrac{\overset{\vdots\;\pi_1(c)}{\Gamma\vdash\Delta,\varphi(c)}}{\Gamma\vdash\Delta,\forall x\varphi(x)}\forall\mathrm{R} \qquad \dfrac{\overset{\vdots\;\pi_2}{\varphi(t),\Theta\vdash\Lambda}}{\forall x\varphi(x),\Theta\vdash\Lambda}\forall\mathrm{L}}{\Gamma,\Theta\vdash\Delta,\Lambda}\,\mathrm{CUT} \;\rightsquigarrow\; \dfrac{\overset{\vdots\;\pi_1(t)}{\Gamma\vdash\Delta,\varphi(t)} \qquad \overset{\vdots\;\pi_2}{\varphi(t),\Theta\vdash\Lambda}}{\Gamma,\Theta\vdash\Delta,\Lambda}\,\mathrm{CUT}$$

5.2.2. *Pushing a cut up the tree.* Now we consider the case of positive depth. The idea is that if the cut formula was already there, we could have done the cut earlier, reducing the depth. Here is an example. The cut formula is χ.

$$\dfrac{\overset{\vdots\;\pi_1}{\Gamma\vdash\Delta,\chi} \qquad \dfrac{\overset{\vdots\;\pi_2}{\chi,\Theta\vdash\Lambda,\varphi} \qquad \overset{\vdots\;\pi_3}{\chi,\Theta\vdash\Lambda,\psi}}{\chi,\Theta\vdash\Lambda,\varphi\wedge\psi}\wedge\mathrm{R}}{\Gamma,\Theta\vdash\Delta,\Lambda,\varphi\wedge\psi}\,\mathrm{CUT} \;\rightsquigarrow\; \dfrac{\dfrac{\overset{\vdots\;\pi_1}{\Gamma\vdash\Delta,\chi} \qquad \overset{\vdots\;\pi_2}{\chi,\Theta\vdash\Lambda,\varphi}}{\Gamma,\Theta\vdash\Delta,\Lambda,\varphi}\,\mathrm{CUT} \qquad \dfrac{\overset{\vdots\;\pi_1}{\Gamma\vdash\Delta,\chi} \qquad \overset{\vdots\;\pi_3}{\chi,\Theta\vdash\Lambda,\psi}}{\Gamma,\Theta\vdash\Delta,\Lambda,\psi}\,\mathrm{CUT}}{\Gamma,\Theta\vdash\Delta,\Lambda,\varphi\wedge\psi}\wedge\mathrm{R}$$

We have again turned one cut into two, but each one is of lower depth, and we can appeal to induction on depth.

This should hopefully give you a feeling of how and why the proof of the cut elimination theorem works. I will leave it at this.

□

As a consequence of the cut elimination theorem, we can prove an important property of intuitionistic logic, showing that it is worthy of the name.

Theorem (Disjunction Property)**.** *If $\vdash \varphi \vee \psi$ is intuitionistically derivable, then either $\vdash \varphi$ or $\vdash \psi$ is intuitionistically derivable.*

Proof. In an intuitionistic cut free proof of $\vdash \varphi \vee \psi$, the last inference can only be one of the $\vee$R rules. □

Likewise,

Theorem (Witness Property)**.** *If $\vdash \exists x\varphi(x)$ is intuitionistically derivable, then for some term t, $\vdash \varphi(t)$ is intuitionistically derivable.*

Proof. In an intuitionistic cut free proof of $\vdash \exists x\varphi(x)$, the last inference can only be the $\exists$R rule. □

What goes wrong in classical logic? In both cases, the last inference could have been a contraction on instances of the sentence that were proved in different ways. This is an example of how structural rules may be less innocent then they appear. We will take up this issue in Herbrand's theorem later.

6. Interpolation

Another application of the cut elimination theorem is to give proof-theoretic proofs of various interpolation theorems. In such a proof, the interpolant is assembled in a constructive process from a cut free proof. We have the following version of Craig's interpolation theorem.

Theorem (Craig Interpolation)**.** *Suppose that $\varphi \vdash \psi$ is derivable. Then there is a θ such that the only nonlogical symbols occurring in θ occur both in φ and in ψ, and both $\varphi \vdash \theta$ and $\theta \vdash \psi$ are derivable.*

I will sketch a proof of Lyndon's stronger interpolation theorem.

Theorem (Lyndon Interpolation)**.** *Suppose that $\varphi \vdash \psi$ is derivable. Then there is a θ such that any relation occurring in a positive (negative) context in θ already occurs in a positve (negative) context in both φ and ψ, and both $\varphi \vdash \theta$ and $\theta \vdash \psi$ are derivable.*

This statement will be deduced as a special case of the following more general statement which behaves better as an inductive hypothesis.

Lemma. *Suppose $\Gamma \vdash \Delta$ is derivable, and $(\Gamma_1, \Gamma_2), (\Delta_1, \Delta_2)$ are respectively partitions of Γ, Δ. Then there is a θ such that $\Gamma_1 \vdash \theta, \Delta_1$ and $\Gamma_2, \theta \vdash \Delta_2$ are derivable, and any relation occurring in a positive (negative) in θ already occurs in a positive (negative) context in both $\Gamma_1 \cup \neg\Delta_1$ and $\neg\Gamma_2 \cup \Delta_2$.*

Proof sketch. This is a proof by induction on the length of a cut free proof of $\Gamma \vdash \Delta$. In other words, we argue by cases on the last inference of such a proof. I will do a few cases.

For a base case, if $\Gamma \vdash \Delta$ is an axiom $\varphi \vdash \varphi$, we argue by cases on the partition:

$$(\varphi, \emptyset),\ (\emptyset, \varphi) \rightsquigarrow \theta = \varphi$$

$$(\varphi, \emptyset),\ (\varphi, \emptyset) \rightsquigarrow \theta = \bot$$

$$(\emptyset, \varphi),\ (\emptyset, \varphi) \rightsquigarrow \theta = \top$$

$$(\emptyset, \varphi),\ (\varphi, \emptyset) \rightsquigarrow \theta = \neg\varphi$$

Suppose now the last inference was a $\wedge$R.

$$\dfrac{\begin{array}{c}\vdots\ \pi_1 \\ \Gamma \vdash \Delta, \varphi\end{array} \qquad \begin{array}{c}\vdots\ \pi_2 \\ \Gamma \vdash \Delta, \psi\end{array}}{\Gamma \vdash \Delta, \varphi \wedge \psi}\ \wedge\mathrm{R}$$

For any partition of the endsequent of this proof, we have a corresponding partition of the endsequents of π_1 and π_2. For instance, given a partition $(\Gamma_1, \Gamma_2), (\Delta_1, \Delta_2, \varphi \wedge \psi)$ we have partitions

$$(\Gamma_1, \Gamma_2), (\Delta_1, \Delta_2 + \varphi) \rightsquigarrow \lambda$$

$$(\Gamma_1, \Gamma_2), (\Delta_1, \Delta_2 + \psi) \rightsquigarrow \varrho$$

In this case our interpolant is $\theta = \lambda \wedge \varrho$. If $\varphi \wedge \psi$ is instead grouped with Δ_1, our interpolant is $\lambda \vee \varrho$. Here the following sequents are derivable:

$$\Gamma_1 \vdash \Delta_1, \varphi, \lambda$$

$$\Gamma_1 \vdash \Delta_1, \psi, \varrho$$

$$\Gamma_2, \lambda \vdash \Delta_2$$

$$\Gamma_2, \varrho \vdash \Delta_2$$

We can infer $\Gamma_1 \vdash \Delta_1, \varphi \wedge \psi, \lambda \vee \varrho$ from the first two, and $\Gamma_2, \lambda \vee \varrho \vdash \Delta_2$ from the last two.

For a case that demonstrates the necessity of the partitions, suppose the last inference was an instance of $\neg$R

$$\dfrac{\begin{array}{c}\vdots\ \pi \\ \varphi, \Gamma \vdash \Delta\end{array}}{\Gamma \vdash \Delta, \neg\varphi}\ \neg\mathrm{R}$$

Consider for instance a partition of the form (Γ_1, Γ_2), $(\Delta_1, \Delta_2 + \neg\varphi)$. For this, we consider the partition $(\Gamma_1, \Gamma_2 + \varphi)$, (Δ_1, Δ_2) of the endsequent of π, and appeal to induction, yielding an interpolant θ such that the following sequents are derivable:

$$\Gamma_1 \vdash \Delta_1, \theta$$

$$\Gamma_2, \varphi, \theta \vdash \Delta_2$$

We can then infer $\Gamma_2, \theta \vdash \Delta_2, \neg\varphi$ from the second one, so θ works as an interpolant for the partition of $\Gamma \vdash \Delta, \neg\varphi$.

I here record one extra tricky case, which may not be presented. Suppose the last inference is an instance of →L:

$$\dfrac{\begin{array}{c}\vdots\,\pi_1\\ \Gamma \vdash \Delta, \varphi\end{array} \qquad \begin{array}{c}\vdots\,\pi_2\\ \psi, \Theta \vdash \Lambda\end{array}}{\varphi \to \psi, \Gamma, \Theta \vdash \Delta, \Lambda} \to\mathrm{L}$$

Suppose also that we are given a partition of the form:

$$(\Gamma_1 + \Theta_1 + \varphi \to \psi, \Gamma_2 + \Theta_2), (\Delta_1 + \Lambda_1, \Delta_2 + \Lambda_2)$$

For this, we consider the following partitions:

$$(\Gamma_2, \Gamma_1), (\Delta_2, \Delta_1 + \varphi) \rightsquigarrow \kappa$$

$$(\Theta_1 + \psi, \Theta_2), (\Lambda_1, \Lambda_2) \rightsquigarrow \eta$$

Thus the following sequents are derivable

$$\Gamma_2 \vdash \Delta_2, \kappa$$

$$\kappa, \Gamma_1 \vdash \Delta_1, \varphi$$

$$\psi, \Theta_1 \vdash \Lambda_1, \eta$$

$$\eta, \Theta_2 \vdash \Lambda_2$$

We will show that $\kappa \to \eta$ works as an interpolant. In other words,

$$\Gamma_1, \Theta_1, \varphi \to \psi \vdash \Delta_1, \Lambda_1, \kappa \to \eta$$

$$\kappa \to \eta, \Gamma_2, \Theta_2 \vdash \Delta_2, \Lambda_2$$

are derivable. We derive these below:

$$\dfrac{\dfrac{\kappa, \Gamma_1 \vdash \Delta_1, \varphi \qquad \psi, \Theta_1 \vdash \Lambda_1, \eta}{\Gamma_1, \Theta_1, \varphi \to \psi, \kappa \vdash \Delta_1, \Lambda_1, \eta} \to\mathrm{L}}{\Gamma_1, \Theta_1, \varphi \to \psi \vdash \Delta_1, \Lambda_1, \kappa \to \eta} \to\mathrm{R} \qquad \dfrac{\Gamma_2 \vdash \Delta_2, \kappa \qquad \eta, \Theta_2 \vdash \Lambda_2}{\kappa \to \eta, \Gamma_2, \Theta_2 \vdash \Delta_2, \Lambda_2} \to\mathrm{L}$$

This is a far from comprehensive treatment, but hopefully it gives you a feeling of how the proof method works.

□

7. Herbrand's Theorem

As mentioned previously, the witness property does not hold for classical logic. In certain settings, however, it holds "up to a finite amount of ambiguity." One way of capturing this is the following weak form of Herbrand's theorem.

Theorem (Herbrand's Theorem: weak form)**.** *Let T be set of universal axioms, and suppose that*

$$\vdash \forall \overline{x} \exists y_1 \ldots \exists y_k \theta(\overline{x}, y_1, \ldots, y_k)$$

is derivable using axioms from T, where θ is quantifier free. Then there are terms $\{\tau_{i,j}(\overline{x})\}_{i \leq k, j \leq m}$ such that

$$\vdash \bigvee_{j \leq m} \theta(\overline{x}, \tau_{1,j}(\overline{x}), \ldots, \tau_{k,j}(\overline{x}))$$

is derivable from axioms in T.

This is weak insofar as it only applies to $\forall\exists$ consequences of $\forall$ theories. Herbrand actually proved a stronger theorem that applies to arbitrary formulas in prenex normal form, which would take longer to state. In any case, this is what we will prove.

We may remove the outermost universal quantifiers of everything, so that T consists of quantifier free formulas and

$$\vdash \exists y_1 \ldots \exists y_k \theta(\overline{x}, y_1, \ldots, y_k)$$

is derivable from axioms in T. By the free cut elimination theorem I mentioned earlier, this has a proof all of whose cuts are on formulas which are axioms in T. In particular, all cuts are on quantifier free formulas.

What can happen on the right side of this proof? Well, we can apply the $\exists$R and CR rules. Undoing these, we should have the conclusion of the theorem.

To do this more formally, we will as usual argue by induction on the size of the proof. In particular, we will deduce the theorem as a special case of the following lemma.

Lemma. *Suppose there is a free-cut-free proof (from T) of*

$$\Gamma \vdash \Delta, \Lambda$$

such that Γ and Δ consist of quantifier free formulas, and Λ consists of formulas $\exists \overline{y} \theta(\overline{x}, \overline{y})$. Then for some

$$\Theta = \theta(\overline{x}, \tau_{1,1}(\overline{x}), \ldots, \tau_{k,1}(\overline{x})), \ldots, \theta(\overline{x}, \tau_{1,r}(\overline{x}), \ldots, \tau_{k,r}(\overline{x}))$$

the sequent $\Gamma \vdash \Delta, \Theta$ is derivable.

Proof. We prove this by induction on the size of the proof. For an axiom it is clear, as it must contain only quantifier free formulas. We call a sequent of the form of the endsequent "transparent." Our first claim is that the sequents immediately above the endsequent are also transparent. This follows by analyzing the possible rules that could have produced the endsequent. All rules except cuts have the subformula property, and it is clear that such rules applied backwards produce a transparent sequent. The only remaining case is a cut on a quantifier free formula. Applied backwards, this adds a quantifier free formula to one side on each branch, and each such addition leaves the sequent transparent.

Now that we know the preceding sequents are transparent, the possibilities for the last inference can be divided into cases. The first case we consider is that the formulas operated on by the last inference live in the quantifier free part, Γ and Δ.

$$\dfrac{\begin{array}{c}\vdots\,\pi_1 \\ \Gamma_1 \vdash \Delta_1, \Lambda\end{array} \qquad \begin{array}{c}\vdots\,\pi_2 \\ \Gamma_2 \vdash \Delta_2, \Lambda\end{array}}{\Gamma \vdash \Delta, \Lambda}\ ?$$

Appealing to induction on π_1 and π_2, we have Θ_1, Θ_2 such that $\Gamma_1 \vdash \Delta_1, \Theta_1$ and $\Gamma_2 \vdash \Delta_2, \Theta_2$ are derivable. We can then infer

$$\dfrac{\dfrac{\Gamma_1 \vdash \Delta_1, \Theta_1}{\Gamma_1 \vdash \Delta_1, \Theta_1, \Theta_2}\ \mathrm{WR} \qquad \dfrac{\Gamma_2 \vdash \Delta_2, \Theta_2}{\Gamma_2 \vdash \Delta_2, \Theta_1, \Theta_2}\ \mathrm{WR}}{\Gamma \vdash \Delta, \Theta_1, \Theta_2}\ ?$$

We now consider the case that the last inference operates on Λ, in other words, the output of the inference is in Λ. By the structure of transparent sequents, the only possibilities are CR and $\exists$R. In the CR case, we have

$$\dfrac{\begin{array}{c}\vdots\,\pi \\ \Gamma \vdash \Delta, \Lambda, \varphi, \varphi\end{array}}{\Gamma \vdash \Delta, \Lambda, \varphi}\ \mathrm{CR}$$

Appealing to induction on π tells us that $\Gamma \vdash \Delta, \Theta$ is derivable for some Θ, which is what we need.

The case where something happens is the $\exists$R case. In this case, we have

$$\dfrac{\begin{array}{c}\vdots\,\pi \\ \Gamma \vdash \Delta, \Lambda, \theta(\overline{x}, \tau_1(\overline{x}), \ldots, \tau_k(\overline{x}))\end{array}}{\Gamma \vdash \Delta, \Lambda, \exists \overline{y} \theta(\overline{x}, \overline{y})}\ \exists\mathrm{R}$$

Appealing to induction,

$$\Gamma \vdash \Delta, \Theta, \theta(\overline{x}, \tau_1(\overline{x}), \ldots, \tau_k(\overline{x}))$$

is derivable. We can then define

$$\Theta' = \Theta, \theta(\overline{x}, \tau_1(\overline{x}), \ldots, \tau_k(\overline{x}))$$

so $\Gamma \vdash \Delta, \Theta'$ is derivable.

$\square$