# TRANSLATIONS BETWEEN LINEAR AND TREE NATURAL DEDUCTION SYSTEMS FOR RELEVANT LOGICS

### SHAWN STANDEFER

## The University of Melbourne

**Abstract.** Anderson and Belnap presented indexed Fitch-style natural deduction systems for the relevant logics R, E, and T. This work was extended by Brady to cover a range of relevant logics. In this paper I present indexed tree natural deduction systems for the Anderson–Belnap– Brady systems and show how to translate proofs in one format into proofs in the other, which establishes the adequacy of the tree systems.

The earliest relevant logics were put forward in the form of Hilbert-style axiom systems.<sup>1</sup> Alan Ross Anderson and Nuel Belnap provided Fitch-style cascade natural deduction systems for what they regarded as the central relevant logics, the logics R, E, and T.<sup>2</sup> The idea behind the systems was that when one introduces a conditional, the antecedent should be relevant to the consequent. The usual conditional rules are the following.

$$\begin{vmatrix} A & & \\ A \to B & \\ B & \to E & \\ A \to B & \to I \end{vmatrix} \xrightarrow{A \to B} A \to A$$

In the  $\rightarrow$ I rule, one only needs to reach the conclusion *B* under the assumption of *A*. Standard cascade systems have a rule of reiteration that permit one to copy sentences from superproofs to subproofs, permitting the following derivation.



Received: November 9, 2016.

<sup>2020</sup> Mathematics Subject Classification: 03F52, 03B47.

Keywords and phrases: relevant logics, natural deduction, translations.

<sup>&</sup>lt;sup>1</sup> See Moh (1950), Church (1953), and Došen (1992). For overviews of the history and development of relevant logics, see Dunn & Restall (2002) and Bimbó (2006).

<sup>&</sup>lt;sup>2</sup> See Anderson & Belnap (1975) and Anderson, Belnap, & Dunn (1992) for the natural deduction systems. See Pelletier (1999) for an overview of the development of different types of natural deduction systems.

Jaśkowski (1934) presented a natural deduction system that is similar to that of Fitch. Since Jaśkowski's work predates Fitch's, Fitch-style systems should perhaps be called "Jaśkowski–Fitch-style systems" or "JF-systems." I will use the term "cascade system" for this style of linear natural deduction system, to have a term that does not involve proper names.

#### SHAWN STANDEFER

To prevent this sort of derivation, Anderson and Belnap required that for a conditional to be introduced, the antecedent had to be used to obtain the consequent. They employed indices to track use, in this sense. The indices are sets of natural numbers. Hypotheses have singleton indices  $\{k\}$ , where k is the depth of nesting of subproofs at that hypothesis. The Anderson–Belnap rules for the conditional are the following.

$$\begin{array}{c|c} A_a & & \\ A \to B_b & \\ B_{a \cup b} & \to \mathbf{E} \end{array} \qquad \begin{array}{c|c} A_{\{k\}} & \text{hyp} \\ \hline \vdots \\ B_a \\ A \to B_{a-\{k\}} & \to \mathbf{I} \end{array}$$

In  $\rightarrow$ I, it is required that  $k \in a$ . This requirement prevents the derivation of (K) axiom, sometimes called the weakening axiom or the law of irrelevance,  $A \rightarrow (B \rightarrow A)$ , above, as the index on A will be {1}, whereas B will have {2}. These rules are the conditional rules for the logic R. To obtain the logics E and T, Anderson and Belnap place restrictions on either the reit rule or  $\rightarrow$ E. Indexed cascade systems can be extended from the arrow fragment to the vocabulary { $\rightarrow$ ,  $\sim$ , &,  $\lor$ } without difficulty. Brady (1984) extended the Anderson–Belnap systems to cover a wide range of relevant logics without any restriction on reit.

Prawitz (1965, 81ff.) gave a tree natural deduction system for a relevant logic. His system did not use indices. Rather, he used side conditions on the rule  $(\rightarrow I)$ . First, he required that each instance of the rule discharge at least one occurrence of an assumption. This no-vacuous discharge policy was sufficient to rule out the derivation of  $A \rightarrow (B \rightarrow A)$ , but the addition of the & and  $\lor$  created some difficulties.

To accommodate the truth-functional connectives, Prawitz imposed a side condition on derivations, in addition to the condition banning vacuous discharge for conditionals. The condition uses the concept of a sentence occurrence *B* originating from a sentence occurrence *A*. *B* originates from *A* just in case there is a sequence of sentence occurrences  $C_1, ..., C_n$ , where  $C_1 = A$ ,  $C_n = B$ , *A* is not discharged by  $(\vee E)$ , and  $C_{i+1}$ is either immediately below  $C_i$  or  $C_{i+1}$  is an assumption discharged by an application of  $(\vee E)$ . The side condition is on applications of  $(\rightarrow I)$  that are below applications of (& I) or  $(\vee E)$ . Let such an application of  $(\Rightarrow I)$  have *B* as its premise and discharge occurrences of *A*. If one premise of (& I), or one minor premise of  $(\vee E)$ , originates from *A* in the subderivation determined by *B*, then the other premise, respectively minor premise, must also originate from an occurrence of A that is also discharged by that application of  $(\rightarrow I)$ .<sup>3</sup>

Prawitz's system has the benefit of using the familiar tree natural deduction rules for & and  $\vee$ .<sup>4</sup> The system normalizes, and it permits a derivation of distribution of conjunction over disjunction:  $A \& (B \vee C) \rightarrow (A \& B) \vee (A \& C)$ .<sup>5</sup> It trades indices for the side condition above. The downside, depending on one's view, is that the logic is not that of R, but rather the positive logic obtained from Urquhart's semilattice semantics.<sup>6</sup> This is a proper strengthening of positive R, as it permits derivation of  $(A \rightarrow B \vee C) \& (C \rightarrow D) \rightarrow (A \rightarrow B \vee D)$ , which is invalid in R. The derivation

$$\frac{(A \to B \lor C) \& (B \to C)}{\underline{A \to B \lor C}} \stackrel{(\&E)}{(\&E)} \qquad \underline{A} \stackrel{(\to E)}{(\to E)} \qquad \frac{(A \to B \lor C) \& (B \to C)}{\underline{B \to C}} \stackrel{(\&E)}{(\&E)} \qquad \underline{B} \stackrel{(1)}{(\to E)} \stackrel{(E)}{(\to E)} \stackrel{(C)}{(\to E)} \stackrel{(C)}$$

can be extended to a derivation of the R-invalid formula by two applications of  $(\rightarrow I)$ . The interested reader can check that Prawitz's side condition holds for this derivation.

Since the publication of Prawitz's work, there appears to have been comparatively little interest in developing tree natural deduction systems for relevant logics.<sup>7</sup> I will briefly mention a few exceptions. First, there has been some interest in tree systems in connection with Urquhart's semilattice semantics. Giambrone & Urquhart (1987) and Meyer *et al.* (1988) gave systems for the positive logics of the semilattice semantics. These systems use indices rather than Prawitz's side condition on  $(\rightarrow I)$ .<sup>8</sup> Second, Slaney (1990), Read (1988), and Restall (2000) present tree natural deduction systems in which nodes are sequents. These approaches cover a large range of relevant logics, and include structural rules operating on the left-hand side of the sequents. I will discuss these systems more in the final section. Finally, Francez (2014) investigates an index-free, signed tree natural deduction system for a positive fragment of R.

In this paper, I will provide indexed tree natural deduction systems for a range of relevant logics ( $\S$ 2). I will first present the cascade systems developed by Brady ( $\S$ 1). I will show the trees sound and complete for the target logics by showing how to translate

<sup>&</sup>lt;sup>3</sup> Prawitz (1965, p. 85). The full side condition covers the classical absurdity rule and includes  $(\exists E)$  in its statement, although those are omitted here.

<sup>&</sup>lt;sup>4</sup> Throughout, & and  $\lor$  bind more tightly than  $\rightarrow$ .

<sup>&</sup>lt;sup>5</sup> In fact, the usual classical derivation will work.

<sup>&</sup>lt;sup>6</sup> See Charlwood (1981). Dunn & Restall (2002) incorrectly state that Prawitz's system is R minus distribution. See Urquhart (1972) for more details on the semilattice semantics and some of its variations, and Humberstone (1988) for a modification to bring the semilattice semantics in line with the positive logic of R.

<sup>&</sup>lt;sup>7</sup> Here I mean relevant logics in the tradition of Anderson, Belnap, Dunn, Meyer, Sylvan, and Plumwood. Tennant (2012, 2015, 2017) provides tree proof systems for core logic, which takes a distinct approach to relevance.

<sup>&</sup>lt;sup>8</sup> This is not to say that no side conditions used. There are some, although they are easier to check, dealing just with the indices.

cascade proofs into tree proofs (\$3), and vice versa (\$4).<sup>9</sup> I will conclude by indicating some possible payoffs of these new systems (\$5).

**§1. Cascade systems.** Anderson and Belnap present cascade systems for the relevant logics R, E, and T, which are relatively strong logics. There are a host of well known, weaker relevant logics.<sup>10</sup> Brady (1984) provided cascade systems for many of these weaker systems.<sup>11</sup> Before getting into the cascade systems, it will be useful to lay out Hilbert-style axioms for the positive logics. I will return to negation later in this section.

The axiom systems are for logics as sets of theorems.<sup>12</sup> All of the logics I will look at have the following axioms and rules.

- (A1)  $A \to A$ (A2)  $A \& B \to A$ (A3)  $A \& B \to B$ (A4)  $(A \to B) \& (A \to C) \to (A \to B \& C)$ (B1)  $A . A \to B \Rightarrow B$
- $(A5) A \rightarrow A \lor B \qquad (R2) A, B \Rightarrow A \& B$
- (A6)  $B \to A \lor B$  (R3)  $(A \to B).(C \to D) \Rightarrow (B \to C) \to$
- (A7)  $(A \to C) \& (B \to C) \to (A \lor B \to C)$

The logics are distinguished by the adoption of some or all of the following axioms.

 $(A \rightarrow D)$ 

 $(C1) (A \to B) \& (B \to C) \to (A \to C)$  $(C2) (A \to B) \to ((B \to C) \to (A \to C))$  $(C3) (A \to B) \to ((C \to A) \to (C \to B))$  $(C4) (A \to (A \to B)) \to (A \to B)$  $(C5) (A \to (B \to C)) \to (B \to (A \to C))$ 

Different positive logics are obtained by adopting some or all of these axioms.<sup>13</sup>

| B A1-A8, R1-R3 | T TJ + C4    |
|----------------|--------------|
| DJ B + C1      | RW TW $+$ C5 |
| TW $B + C2-3$  | R RW + C4    |
| $TJ\;TW+C1$    |              |

<sup>&</sup>lt;sup>9</sup> Von Plato (2017) provides translations between cascade, there called "linear," and tree natural deduction systems for intuitionistic and classical logic. The cascade-to-tree translation is similar to the one adopted here. The tree-to-cascade translation is substantially different. The cascade-to-tree translation of Standefer (2019) is roughly the same as the one used here, but the tree-to-cascade translation used here is substantially different from that of Standefer (2019). The present translation is an improvement in not requiring the shuffling around and subsequent identification of assumptions, instead using the extra information contained in the indices to identify assumptions at the outset.

<sup>&</sup>lt;sup>10</sup> Many of these weaker logics are studied by Routley *et al.* (1982), among other sources.

<sup>&</sup>lt;sup>11</sup> Brady (1984) also provided systems for quantificational logics and rule strengthenings of these logics. I will not get into those systems here.

<sup>&</sup>lt;sup>12</sup> They are in the framework FMLA of Humberstone (2011, 103ff.).

<sup>&</sup>lt;sup>13</sup> A more careful nomenclature for the logics would include a superscripted '+' to indicate that positive logics, i.e. negation-free logics, are under consideration. Since these are the only logics under consideration for most of this paper, I will leave the superscript off.

With the target logics in view, we can now turn to the cascade systems.

The cascade systems have an intuitive graphical presentation. For the development of this paper, it will be useful to present them in a nongraphical form, which will permit precise definitions.<sup>14</sup> To settle some notation, X and Y will be used for sequences of formulas, X, A will be used for the concatenation of X and  $\langle A \rangle$ ,  $X^{\frown} \langle A \rangle$ , and X, Y will be used for the concatenation  $X \cap Y$ . The lower case letters a, b, c, d, and e are used for sets of natural numbers and lower case letters from later in the alphabet, i, j, k, etc., as well as  $\lambda$ , are used for natural numbers.

DEFINITION 1. A sequent,  $A_{\{k_1\}}^1, \dots, A_{\{k_n\}}^n \succ_{\mathfrak{C}} B_b$ , is a pair of a sequence,  $\langle A_{\{k_1\}}^1, \dots, A_{\{k_n\}}^n \rangle$ , of indexed formulas and an indexed formula  $B_b$ .<sup>15</sup> If n = 0, then the sequence is  $\langle \rangle$ .

The length  $\mathfrak{L}(X)$  of a sequence X is defined recursively as  $\mathfrak{L}(\langle \rangle) = 0$  and  $\mathfrak{L}(X \cap \langle A \rangle) =$  $\mathfrak{L}(X) + 1.$ 

The depth of a sequent  $s = X \succ_{\mathfrak{C}} A$ ,  $\mathfrak{D}(s)$ , is  $\mathfrak{L}(X)$ .

An S-sequence is a sequence of sequents.<sup>16</sup>

Let  $\Sigma$  be an S-sequence,  $\langle s_1, \dots, s_k \rangle$ . The depth of the *n*th member of  $\Sigma$ ,  $\mathfrak{D}(n)$ , is  $\mathfrak{D}(s_n).$ 

DEFINITION 2. An S-sequence,  $(s_1, ..., s_n)$ , is nested iff for all m < n,  $s_m$  and  $s_{m+1}$ have one of the following forms: either

- s<sub>m</sub> is X≻<sub>C</sub>A<sub>a</sub> and s<sub>m+1</sub> is X, B<sub>{k</sub>≻<sub>C</sub>B<sub>{k</sub>}</sub>.
  s<sub>m</sub> is X≻<sub>C</sub>A<sub>a</sub> and s<sub>m+1</sub> is X≻<sub>C</sub>B<sub>b</sub>, or
- $s_m \text{ is } X, A_{\{k\}} \succ_{\mathfrak{C}} B_b \text{ and } s_{m+1} \text{ is } X \succ_{\mathfrak{C}} A \to B_{b-\{k\}},$

provided  $n \ge 2$ . If n = 1, then  $s_1$  is of the form  $B_{\{k\}} \succ_{\mathfrak{C}} B_{\{k\}}$ .

The third case in the definition deserves comment. The third case is the one in which a hypothesis is no longer available. For the purposes of this paper, it is sufficient to consider proofs in which a hypothesis is discharged via a conditional introduction rule. This is not the most general form of the definition, but it will work for present purposes. The addition of quantifiers or a modal operator, or the use of a different disjunction elimination rule or another rule that discharges assumptions, would require altering this definition.

For many purposes, it is best to picture the S-sequences written out vertically, rather than horizontally. I will refer to the individual members of an S-sequent as lines. The indexed formula  $A_a$  in a line  $X \succ_{\mathfrak{C}} A_a$  is the formula of the line. For reasons that will become clear soon, I will call a sequent of the form  $X, A_{\{k\}} \succ_{\mathfrak{C}} A_{\{k\}}$ , where no  $k' \ge k$  occurs in X a hyp step. I will also use the notation  $n \in [i, j]$  for  $i \le n \le j$ , with parentheses replacing one or both brackets to mean the corresponding ordering is strict, e.g.  $n \in (i, j]$  is  $i < n \le j$ .

DEFINITION 3 (Subproof, availability). Let  $\Sigma$  be a nested S-sequence. A subproof  $\sigma$  of  $\Sigma$  is the subsequence  $\langle s_i, \dots, s_i \rangle$ , where  $1 \le i \le j \le \mathfrak{L}(\Sigma)$  such that

<sup>14</sup> Borghuis (1998) presented an alternative precise definition of these sort of natural deduction systems. I would like to thank Valeria de Paiva for the reference.

<sup>15</sup> The superscripts here do the job of indicating possibly distinct sentences, which job is, in other contexts, often done by subscripts.

<sup>&</sup>lt;sup>16</sup> This terminology is based on the terminology of Indrzejczak (2010, chap. 2).

- $s_i$  is  $X, A_{\{k\}} \succ_{\mathfrak{C}} A_{\{k\}}$ ,
- for all  $n \in [i, j]$ ,  $\mathfrak{D}(i) \leq \mathfrak{D}(n)$ ,
- either  $\mathfrak{D}(j+1) = \mathfrak{D}(i) 1$  or  $j = \mathfrak{L}(\Sigma)$ ,

Such a subproof will be denoted  $\sigma_i^i$ .

A subproof  $\sigma_j^i$  is *closed* at line *n* iff j < n, and otherwise  $\sigma_j^i$  is *open* at line *n*. For a closed subproof  $\sigma_i^i$ , *j* is the line that closes the subproof.

The innermost subproof at line n,  $\sigma(n)$ , is the subpool  $\sigma_j^i$ , where  $n \in [i, j]$  and  $\mathfrak{D}(i) = \mathfrak{D}(n)$ .

One subproof  $\sigma$  contains another subproof  $\sigma'$ ,  $\sigma \sqsupset \sigma'$  iff for some  $n, m, j, k, \sigma$  is  $\sigma_m^n$ and  $\sigma'$  is  $\sigma_k^j$ , n < j, and either k < m or both m = k and m is the last line of  $\Sigma$ . The reflexive closure of  $\Box$  is  $\supseteq$ .

A line *n* is available to a line *m*,  $n \succ m$ , iff n < m and either  $\mathfrak{D}(n) = 0$  or  $\sigma(n) \supseteq \sigma(m)$ .

A subproof  $\sigma_j^i$  is available to a line *n* iff either j + 1 = n or  $j + 1 \succ n$ .

These definitions are meant to capture, formally, the usual features of the graphical presentation of cascade proofs. A subproof is a sequence of lines beginning with a hyp step and ending when the formula assumed in the hyp step is no longer an assumption. One subproof  $\sigma$  contains another  $\sigma'$  when  $\sigma'$  begins in the scope of  $\sigma$  and either ends in the scope of  $\sigma$  or both continue till the end of the proof. One line is available to another just in case the former happens earlier, either categorically, i.e. not in a subproof, or with an appropriate subproof containment.

The definition of a closed subproof being available to a line relies on a feature of the definition of nesting: in stepping out of a closed subproof to the the line following the ending of that subproof the depth must decrease by exactly 1. While there are more flexible definitions one could use, it turns out that this will be enough for our purposes.

A proof be a nested S-sequence labeled with rules and justification numbers. The rules will be the following. The conventions on justification numbers will be explained following the definition of a proof.

- **Hyp**  $X, A_{\{k\}} \succ_{\mathfrak{C}} A_{\{k\}}$ , where k > k', for any number k' occurring in an index in X.
- **Reit** From  $X \succ_{\mathfrak{C}} A_a$  to infer  $X, Y \succ_{\mathfrak{C}} A_a$ .
- $\rightarrow \mathbf{E}$  From  $X \succ_{\mathfrak{C}} A \rightarrow B_b$  and  $X \succ_{\mathfrak{C}} A_a$ , to infer  $X \succ_{\mathfrak{C}} B_{a \cup b}$ .
- $\rightarrow$ **I** From a subproof starting with  $X, A_{\{k\}} \succ_{\mathfrak{C}} A_{\{k\}}$  ending with  $X, A_{\{k\}} \succ_{\mathfrak{C}} B_b$ , to infer  $X \succ_{\mathfrak{C}} A \rightarrow B_{b-\{k\}}$ , provided  $k \in b$ .
- &E From  $X \succ_{\mathfrak{C}} A \& B_a$ , to infer  $X \succ_{\mathfrak{C}} A_a$  or  $X \succ_{\mathfrak{C}} B_a$ .
- &I From  $X \succ_{\mathfrak{C}} A_a$  and  $X \succ_{\mathfrak{C}} B_a$ , to infer  $X \succ_{\mathfrak{C}} A \& B_a$ .
- $\forall \mathbf{E} \quad \text{From } X \succ_{\mathfrak{C}} A \to C_b, X \succ_{\mathfrak{C}} B \to C_b, \text{ and } X \succ_{\mathfrak{C}} A \lor B_a, \text{ to infer } X \succ_{\mathfrak{C}} C_{a \cup b}.$
- $\forall \mathbf{I}$  From  $X \succ_{\mathfrak{C}} A_a$  or  $X \succ_{\mathfrak{C}} B_a$ , to infer  $X \succ_{\mathfrak{C}} A \lor B_a$ .
- **Dist** From  $X \succ_{\mathfrak{C}} A \& (B \lor C)_a$ , to infer  $X \succ_{\mathfrak{C}} (A \& B) \lor (A \& C)_a$ .

Call the index  $\{k\}$  on the displayed formula  $A_{\{k\}}$  in the hyp rule the *hyp index*. The hyp rule above is more flexible than Brady's formulation of the rule, a point to which I will return following the definition of a cascade proof.

Now, for the conventions on justification numbers. The rule hyp does not require any justification numbers. A step justified by the rule  $\rightarrow$ I is annotated with the range i - j, where  $\sigma_i^i$  is the subproof indicated in the rule. The justification numbers for

290

the remaining rules are those for lines containing the premise sequents of the rules, in the order presented in the rule statement.

DEFINITION 4. A proof  $\Sigma$  is a nested S-sequence labeled with rules and justification numbers such that

- line 1 is a hyp step with depth 1,
- each member of  $\Sigma$  is the conclusion of one of the rules, which labels it, and whose justification numbers are those of the displayed premise sequents or subproof, and
- all premise sequents or subproofs cited in justification numbers for a line λ are available to λ.

It will be useful in the translations to treat the lines of a proof as including the rules and justifications, which I will do.

I can now proceed to Brady's side conditions. These are conditions on the  $\rightarrow E$  and  $\lor E$  rules, with max $(\emptyset) =_{Df} - 1$ .

 $\begin{array}{ll} \mathsf{B} & \text{ If } b \neq \emptyset, \text{ then } a = \{k\} \text{ such that } \max(b) < k. \\ \mathsf{DJ} & \text{ If } b \neq \emptyset, \text{ then } a \neq \emptyset, \max(b) < \max(a), \text{ and } a - \{\max(a)\} \in \{b, \emptyset\}. \\ \mathsf{TW} & \text{ If } b \neq \emptyset, \text{ then } a \neq \emptyset, \max(b) < \max(a), \text{ and } a \cap b = \emptyset. \\ \mathsf{TJ} & \text{ If } b \neq \emptyset, \text{ then } a \neq \emptyset, \max(b) < \max(a), \text{ and either } a \cap b = \emptyset \text{ or } a - \{\max(a)\} = b. \\ \mathsf{T} & \text{ If } b \neq \emptyset, \text{ then } a \neq \emptyset, \max(b) \leq \max(a). \\ \mathsf{RW} & a \cap b = \emptyset. \\ \mathsf{R} & \text{ No restriction.} \end{array}$ 

Brady requires the hyp index to be the rank of the formula, defined as the depth of the line of the hyp rule. The additional flexibility of the hyp rule adopted here does not lead to new theorems being provable. Any proof using this hyp rule can be transformed into a proof using Brady's hyp rule by shifting hyp indices that are greater than the rank to the rank throughout the proof and percolating the results through the rules. That the result satisfies the conditions is a consequence of the following lemma, for which it will be useful to employ a definition.

**DEFINITION 5.** A partial function  $f: \omega \rightarrow \omega$  is

- *monotonic* iff n < m implies  $f(n) \le f(m)$ , for all n, m in the domain of f;
- *injective* iff f(n) = f(m) implies n = m, for all n, m in the domain of f.

For a set of natural numbers b contained in the domain of a partial function f,  $f[b] = \{f(n) : n \in b\}.$ 

LEMMA 1. Suppose a monotonic, injective partial function  $f: \omega \rightarrow \omega$  is defined on a finite set  $a \cup b$ , where  $a, b \subseteq \omega$ . Then the following hold.

1. If  $a \neq \emptyset$  then  $f[a] \neq \emptyset$ .

- 2. If a is a singleton, then f[a] is a singleton.
- 3. If  $\max(a) < \max(b)$ , then  $\max(f[a]) < \max(f[b])$ .
- 4. If  $a \cap b = \emptyset$ , then  $f[a] \cap f[b] = \emptyset$ .
- 5. If  $a \max(a) = b$ , then  $f[a] \max(f[a]) = f[b]$ .

**PROOF.** Suppose a monotonic, injective partial function *f* is defined on a finite set  $a \cup b$ , where  $a, b \subseteq \omega$ .

Both 1 and 2 are immediate.

For 3, suppose  $\max(a) < \max(b)$ . Since f is monotonic,  $\max(f[a]) \le \max(f[b])$ . By the injective condition on f,  $\max(f[a]) \ne \max(f[b])$ .

For 4, suppose  $a \cap b = \emptyset$ , but  $f[a] \cap f[b] \neq \emptyset$ . Then there is some  $n \in f[a] \cap f[b]$ , and so  $n \in f[a]$  and  $n \in f[b]$ . But then there are *m* and *k* such that  $m \in a, k \in b$  and f(m) = f(k) = n. By the injective condition, m = k, so  $a \cap b \neq \emptyset$ , contradicting the assumption.

For 5, suppose  $a - \max(a) = b$ . Suppose  $f[a] - \max(f[a]) \neq f[b]$ . There are two cases. Suppose there is  $n \in f[a] - \max(f[a])$  but  $n \notin f[b]$ . Then for  $m = f(n), m \in a - \max(a)$  but  $m \notin b$ , contradicting the assumption. Suppose there is  $n \notin f[a] - \max(f[a])$  and  $n \in f[b]$ . Then for  $m = f(n), m \notin a - \max(a)$  but  $m \in b$ , contradicting the assumption.

Shifting indices in a sequent so that hyp indices match the rank of a step defines a monotonic, injective partial function on indices in a sequent. The indices on the steps in the original cascade proof satisfy the required conditions. This is sufficient to apply lemma 1, as appropriate for the target logic, for the corresponding steps in the proof with shifted indices.

I will focus on the positive logics for this paper. The addition of negation adds no new challenges, just a proliferation of straightforward cases. The techniques of this paper can easily be extended to cover negation rules. The issue, rather, is that the addition of negation would somewhat clutter the presentation. The full logics, along with other logics, such as DW and DK, that differ from ones above only in their negation axioms can be accommodated by the techniques below.

Next, I need to introduce some notation. I will use the notation  $A_{a_1}^1, ..., A_{a_n}^n \vdash_{\mathfrak{C}} B_b$  to indicate that there is a proof whose final line is  $A_{a_1}^1, ..., A_{a_n}^n \succ_{\mathfrak{C}} B_b$ , with appropriate rule and justification annotations.<sup>17</sup> I will say that  $\Sigma$  is a proof of  $A_{a_1}^1, ..., A_{a_n}^n \vdash_{\mathfrak{C}} B_b$  just in case  $\Sigma$  witnesses the claim that  $A_{a_1}^1, ..., A_{a_n}^n \vdash_{\mathfrak{C}} B_b$ .

Let us proceed to the tree natural deduction proof systems.

§2. Trees. The tree natural deduction proof systems I will give use indices, in much the same way as the Anderson–Belnap–Brady systems of the previous section. The technical development of the tree systems will follow that of Prawitz (1965), which will largely be suppressed here. There is a rule of assumption,  $A_{\{k\}}$ , whose index is a singleton  $\{k\}$ , where k is an natural number. The other rules are given in Figure 1.<sup>18</sup>

The rule  $(\rightarrow I)$  has three conditions: (i)  $k \in a$ , (ii) all occurrences of open assumptions of  $A_{\{k\}}$  in the subtree ending with the premise of the rule are discharged, and (iii) no other assumption in the subtree ending with the premise of the rule has an index  $\{k'\}$  with k' > k. The third condition can be dropped in the cases of R and RW, but its inclusion does not cause any problems. The condition is needed in the other cases.

Condition (ii) can be motivated by reflection on some ideas of Leivant (1979). In the tree systems, the indices on assumptions mark out assumption classes, in the sense

<sup>&</sup>lt;sup>17</sup> Note that since each hyp index  $a_i$  is  $\{k_i\}, k_i < k_{i+1}$ , for  $i \in [1, n)$ .

<sup>&</sup>lt;sup>18</sup> I will use  $\Pi, \Pi_1, ...$  to include the step listed immediately above the rule line in the tree proofs.

Fig. 1. Rules of the tree systems.

of Leivant (1979). An *assumption class* is a set of occurrences of an indexed formula in a proof. Leivant was working with unindexed formulas and so used Prawitz's notion of a discharge function to distinguish members of a particular assumption class. In this setting, the indices on assumption formulas mark out the assumption classes and so carry the information needed for a discharge function for a proof.

Condition (iii) may seem slightly *ad hoc*, so a few words motivating it are in order. The condition is actually in force in cascade proofs, although it is hidden in plain sight. In graphical form, the rule for introducing the conditional is the following.

$$n \qquad \begin{vmatrix} A_{\{k\}} & \text{hyp} \\ \hline \vdots \\ m & B_a \\ p & A \to B_{a-\{k\}} & \to \mathbf{I} \ n-m \end{vmatrix}$$

The hypothesis discharged,  $A_{\{k\}}$ , is the assumption of the innermost subproof at line *m*. Since the hyp index for line *n* is *k*, and the subproof started on line *n* is the innermost, there is no k' > k such that an undischarged hypothesis has index  $\{k'\}$ . There is no need to postulate an analog of (iii) for cascade proofs because the subproof structure builds it in.<sup>19</sup>

Finally, we must supply global constraints on tree proofs for the index discipline we adopt. First, assumptions have indices that are singletons,  $\{k\}$ , where k is a natural number. No open assumptions of distinct formulas can have the same index. An open assumption may have the same index as a discharged assumption, depending on the construction of the proof. Different occurrences of a single formula as open assumptions may have the same index, but they need not.

<sup>&</sup>lt;sup>19</sup> Adjusting cascade systems so that the subproof structure does not force the condition would be a large departure. The result would likely be something similar to Lemmon-style systems. See Lemmon (1965), Slaney (1990), or Indrzejczak (2010, 43ff.) for more on Lemmon-style systems.

These are the rules and conditions common to all of the tree systems. To obtain the different logics of the previous section, one imposes Brady's conditions on the  $(\lor E)$  and  $(\rightarrow E)$  rules.

As in the previous section, I will define some notation for being a proof of a formula under certain assumptions. The notation  $A_{a_1}^1, ..., A_{a_n}^n \vdash_{\mathfrak{T}} B_b$  means that there is a tree proof of  $B_b$  whose sequence of open assumption classes is  $A_{a_1}^1, ..., A_{a_n}^n$ . I will say that  $\Pi$ is a proof of  $A_{a_1}^1, ..., A_{a_n}^n \vdash_{\mathfrak{T}} B_b$  just in case  $\Pi$  is a proof witnessing that  $A_{a_1}^1, ..., A_{a_n}^n \vdash_{\mathfrak{T}} B_b$ .

Now, I will move on to the translations.

**§3.** Translations: from cascades to trees. In this section, I will show how to translate cascade proofs to tree proofs, and conversely. I will begin with the translation of cascade proofs to tree proofs. As noted in 1, the justification annotations play an important role in this paper. They enable the definition of trace trees, which are key to this direction of the translation.

Given such a proof, we will construct a tree, called the trace tree. The nodes of a trace tree will have one of three forms:  $A_a$ ,  $[A_{\{k\}}, n]$ , and  $(A_{\{k\}}, n)$ . The latter two are intermediaries that help deal with discharging of assumptions and will be removed once the initial translation is completed. In the translation, assumptions are translated as  $[A_{\{k\}}, n]$ , and  $(A_{\{k\}}, n)$  marks which occurrences of  $[A_{\{k\}}, n]$  will be replaced by discharged assumptions, with the other occurrences of the bracketed pairs replaced by undischarged assumptions. The trace tree  $t(\lambda)$  of a line  $\lambda$  of a proof  $\Sigma$  is defined inductively as follows.

Case: Hyp. Line  $\lambda$  is  $X, A_{\{k\}} \succ_{\mathfrak{C}} A_{\{k\}}$ , hyp. Then,  $\mathfrak{t}(\lambda)$  is

 $[A_{\{k\}},\lambda]$ 

Case: Reit. Line  $\lambda$  is  $X \succ_{\mathfrak{C}} A_a$ , reit k. Then,  $\mathfrak{t}(\lambda)$  is  $\mathfrak{t}(k)$ . Case:  $\rightarrow I$ . Line  $\lambda$  is  $X \succ_{\mathfrak{C}} A \rightarrow B_{a-\{k\}}$ ,  $\rightarrow I \ i - j$ . Then,  $\mathfrak{t}(\lambda)$  is the following.

$$\frac{(A_{\{k\}},i) \quad \mathfrak{t}(j)}{A \to B_{a-\{k\}}} (\to I)$$

Case: &E,  $\forall$ I, Dist. Suppose  $\rho \in \{$ &E,  $\forall$ I, Dist $\}$ . Line  $\lambda$  is  $X \succ_{\mathfrak{C}} A_a$ ,  $\rho n$ . In that case,  $\mathfrak{t}(\lambda)$  is the following.

$$\frac{\mathfrak{t}(n)}{A_{a}}(\rho)$$

Case:  $\rightarrow E$ , &I. Suppose  $\rho \in \{\rightarrow E, \&I\}$ . Line  $\lambda$  is  $X \succ_{\mathfrak{C}} A_a$ ,  $\rho m, n$ . Then,  $\mathfrak{t}(\lambda)$  is the following.<sup>20</sup>

 $<sup>^{20}</sup>$  Throughout, I will assume that the justifications match the order of the rules presented in this section.

$$\frac{\mathfrak{t}(m)}{A_a} \quad \mathfrak{t}(n)$$

Case:  $\forall E$ . Line  $\lambda$  is  $X \succ_{\mathfrak{C}} C_{a \cup b}$ ,  $\forall E m, n, p$ . In that case,  $\mathfrak{t}(\lambda)$  is the following.

$$\frac{\mathfrak{t}(m) \qquad \mathfrak{t}(n) \qquad \mathfrak{t}(p)}{C_{a\cup b}} (\lor E)$$

The trace tree  $t(\lambda)$  for a line  $\lambda$  of a cascade proof may not be a tree proof. There may be nodes of of the form  $(A_{\{k\}}, i)$  or  $[A_{\{k\}}, i]$ . The tree must be operated upon to remove these nodes. Further, the tree may not satisfy the global constraints. These will be addressed in turn.

To deal with the nonformula nodes, we will construct a sequence of trees. First note that each node  $\alpha$  of the form  $(A_{\{k\}}, i)$  occurs next to a subtree  $\pi_{\alpha}$  ending in the premise  $B_a$  of a  $(\rightarrow I)$  rule. This will be used to define an ordering on nodes  $\alpha$  of the form  $(A_{\{k\}}, i)$ . Say that for nodes  $\alpha$  and  $\beta$ ,  $\alpha < \beta$  iff  $\beta$  occurs in  $\pi_{\alpha}$ .

Let  $\Pi_0$  be the initial tree,  $\mathfrak{t}(\lambda)$ . To form  $\Pi_{j+1}$ , for  $j \ge 0$ , we do the following. Let  $\alpha$  be a <-minimal node in  $\Pi_j$ . In the subtree  $\pi_\alpha$  ending with  $B_a$ , there are nodes of the form  $[A_{\{k\}}, i]$ , where  $A_{\{k\}}$  is discharged by the corresponding application of the rule  $\rightarrow$ I in the cascade proof. Replace all occurrences of  $[A_{\{k\}}, i]$  with  $A_{\{k\}}$ , and mark them as discharged by the  $(\rightarrow I)$  following  $\pi_\alpha$ . Finally, delete the node  $(A_{\{k\}}, i)$  and add (i) to the rule label  $(\rightarrow I)$ . The result is  $\Pi_{j+1}$ . There are fewer nodes of the forms  $(A_{\{k\}}, i)$  in  $\Pi_{j+1}$  than in  $\Pi_j$ , so there must be a  $j \ge 0$  such that  $\Pi_j$  contains no nodes of that form. Call this tree  $\Pi'$ .

The tree  $\Pi'$  may contain nodes of the form  $[A_{\{k\}}, i]$ . Replace each node of the form  $[A_{\{k\}}, i]$  with  $A_{\{k\}}$ . These are the undischarged assumptions of the proof. Call the resulting tree  $\Pi$ . It remains to check that  $\Pi$  satisfies the conditions on the rules and the global conditions on indices.

To begin, there are a two observations concerning  $\Pi$ . First, if  $a \neq \emptyset$ , then an occurrence of indexed formula  $A_a$  in  $\Pi$  has above it, for each  $k \in a$ , an undischarged assumption whose index is  $\{k\}$ . Second, observe that each inference labeled with one of  $(\lor I)$ , (Dist), (&I), and (&E) is an application of that rule.

Next, we show that  $\Pi$  satisfies the restrictions on  $(\rightarrow I)$ . Conditions (i) and (ii) are immediate from the construction. For condition (iii), note that in the  $\rightarrow$ I case of the construction of the trace tree, all hyp steps occurring in  $\sigma_j^i$ , apart from the initial hyp step of  $\sigma_j^i$  have hyp indices greater than that of line *i*, and the construction of  $\Pi$  does not change any indices. Suppose that there is a violation of condition (iii) for some application  $\alpha$  of  $(\rightarrow I)$  such that  $A_{\{k_\alpha\}}$  is discharged but for some  $k > k_\alpha$ ,  $B_{\{k\}}$  is an open assumption of  $\pi_\alpha$ . Suppose the former is the translation of hyp line *n* and the latter of hyp line *m*. There are two cases, n < m and n > m, since n = m is impossible. Suppose n < m. As  $B_{\{k\}}$  is an open assumption in  $\pi_\alpha$  but  $A_{\{k_\alpha\}}$  is discharged by  $\alpha$ , the subproof  $\sigma(m)$  cannot close before  $\sigma(n)$  does. But this violates the conditions on nesting in  $\Sigma$ . If n > m, then the restriction the rule hyp that hyp indices increase with depth is violated. Therefore,  $\alpha$  cannot be a violation of (iii). So, there are no violations of (iii). Next, we argue that  $\Pi$  satisfies the restrictions on  $(\rightarrow E)$  and  $(\lor E)$ .  $\Pi$  satisfies the restrictions because the corresponding steps in  $\Sigma$  did, and the construction of  $\Pi$  does not alter any indices.

All that remains is to check that the global conditions on indices are satisfied. The condition that assumptions have singleton indices holds, as no step in the translation added additional numbers to indices. No open assumptions of distinct formulas in  $\Pi$  have the same index. The construction did not change the indices on the open assumptions of the cascade proof, which were all distinct as required by the hyp rule. As all the conditions are satisfied, I conclude that  $\Pi$  is the desired tree proof.

To translate a cascade proof  $\Sigma$  of  $A_{a_1}^1, ..., A_{a_n}^n \vdash_{\mathfrak{C}} B_b$ , one constructs  $\mathfrak{t}(\mathfrak{L}(\Sigma))$ . One then constructs a series proofs,  $\Pi_0, ..., \Pi_n$ , marks assumptions discharged, removes certain nodes, and finally changes the open assumption nodes from the form  $[A_{\{k\}}, i]$  to  $A_{\{k\}}$ . The resulting proof  $\Pi$  is a tree proof of  $A_{a_{i_1}}^{i_1}, ..., A_{a_{i_m}}^{i_m} \vdash_{\mathfrak{T}} B_b$ . Call the translation just described  $\mathbb{T}$ . So  $\mathbb{T}(\Sigma) = \Pi$ . This gives us the following theorem.

THEOREM 1. Let  $\Sigma$  be a cascade proof of  $A_{a_1}^1, ..., A_{a_n}^n \vdash_{\mathfrak{C}} B_b$ . Then  $\mathbb{T}(\Sigma) = \Pi$  is a tree proof of  $A_{a_{i_1}}^{i_1}, ..., A_{a_{i_m}}^{i_m} \vdash_{\mathfrak{T}} B_b$ .

The proof  $\Pi$  may leave out some steps of  $\Sigma$ , including some open assumptions of  $\Sigma$ . These are steps, and hypotheses, that do not show up in the trace tree of the conclusion. Such steps are, in a sense, otiose for obtaining the conclusion, so it is not, I think, a problem that they are omitted by the translation.

Let us turn to translations in the other direction, from trees to cascades.

**§4.** Translations: from trees to cascades. The translation from tree proofs to cascade proofs uses a technique that I will call *weaving*, which I will define following some auxiliary definitions.

DEFINITION 6. Suppose the cascade proofs  $\Sigma_1, ..., \Sigma_n$  are, respectively, proofs of  $X_i \succ_{\mathfrak{C}} A^i_{a_i}$ . Then the proofs are *compatible* iff for formulas  $B_{\{k_1\}}$  and  $C_{\{k_2\}}$  occurring in the  $X_i$ 's, if  $k_1 = k_2$  then B = C.

The idea for the translation is that subtrees of the tree proof will translate to compatible cascade proofs, which will in turn be woven together. The index discipline on tree proofs will enforce the differentiated condition. To proceed, we will need a definition.

DEFINITION 7. Let  $\Sigma$  be a cascade proof of  $A_{\{k_1\}}^1, ..., A_{\{k_n\}}^n \vdash_{\mathfrak{C}} B_b$  where  $\lambda_i$  and  $\lambda_j$  are the lines for the hyp steps for  $A_{\{k_i\}}^i$  and  $A_{\{k_i\}}^j$ , where j = i + 1 and  $1 \le i < j \le n$ .

The *top* of  $\sigma(\lambda_i)$  with respect to  $\lambda_j$  is the sequence of lines  $\langle s_{\lambda_i}, ..., s_m \rangle$ , where  $\lambda_j = m+1$ .

The *truncated top* of  $\sigma(\lambda_i)$  with respect to  $\lambda_j$  is the sequence of lines  $\langle s_{\lambda_i+1}, ..., s_m \rangle$ , where  $\lambda_j = m + 1$ . If  $\lambda_i = m$ , then the truncated top is  $\langle \rangle$ .

In the top of an open subproof, there may be hyp steps, although these will be for subproofs that close in the top of that subproof. So, the depth of the top of an open subproof may increase before decreasing again to its initial depth.

Suppose one has a pair of compatible cascade proofs,  $\Sigma_1$  of  $A_{a_1}^1, ..., A_{a_n}^n \vdash_{\mathfrak{C}} B_b$  and  $\Sigma_2$  of  $C_{c_1}^1, ..., C_{c_m}^m \vdash_{\mathfrak{C}} D_d$ . Let  $(E_{e_1}^1, p_1), ..., (E_{e_\ell}^\ell, p_\ell)$  be a sequence of pairs of the open

hypotheses of  $\Sigma_1$  and  $\Sigma_2$  together with natural numbers such that  $e_i \leq e_{i+1}$ , for  $1 \leq i < \ell$ and  $p_i$  is 1 or 2, depending on whether  $E_{e_i}^i$  is an open assumption of exactly one of the two proofs or of both. It will be notationally convenient to have a way to refer to the *E* formulas in order of their subscripts. Let *X* be  $\langle A_{a_1}^1, \dots, A_{a_n}^n \rangle$  and let *Y* be  $\langle C_{c_1}^1, \dots, C_{c_m}^m \rangle$ . Then  $\mu(X, Y)$  is  $\langle E_{e_1}^1, \dots, E_{e_p}^\ell \rangle$ .

We construct a sequence of proofs,  $\Sigma^i$ , where  $\Sigma^0 = \langle \rangle$ , and a series of partial functions defined on initial segments of the natural numbers  $f^i : \omega \times \{1,2\} \rightarrow \omega, 0 \le i < \ell$ . The notation  $\sigma_j$  will be used for subproofs from  $\Sigma_j$ , with  $j \in \{1,2\}$ . Given the proof constructed up to stage i,  $\Sigma^i$ , further lines are appended as indicated below. Let Xbe the sequence  $E_{e_1}^1, \dots, E_{e_i}^i$ , and if  $i = 0, X = \langle \rangle$ . Where s(Y) is a sequent with the displayed sequence Y of indexed formulas, s(X) is s(Y), replacing Y with X. There are two basic cases, depending on whether  $p_{i+1} = 1$  or  $p_{i+1} = 2$ .

Case 1: Embed. In this case,  $p_{i+1} = 1$ , so  $E_{e_{i+1}}^{i+1}$  is an assumption of only one proof,  $\Sigma_j$ , and it is the hyp step on line  $\lambda$ ,  $Y, E_{e_{i+1}}^{i+1} \succ_{\mathfrak{C}} E_{e_{i+1}}^{i+1}$ , where Y is a possibly empty sequence of indexed formulas. Append to  $\Sigma^i$  the sequence  $\langle s_r(X) \dots, s_p(X) \rangle$ , where  $\langle s_r(Y) \dots, s_p(Y) \rangle$  is the top of  $\sigma_j(\lambda)$ .

Let  $\lambda_j$  be the greatest line number from  $\Sigma_j$  such that  $f^i(\lambda_j, j)$  is defined.<sup>21</sup> Set  $f^{i+1}(n, j) = f^i(n, j)$ , where  $n \leq \lambda_j$ . For each line  $s_q(Y)$  in the top of  $\sigma_j(\lambda)$ , set f(q, j) = m, where *m* is the line number of the newly added  $s_q(X)$  in  $\Sigma^{i+1}$ . Finally, each justification number *n* in  $\Sigma^{i+1}$  is replaced by f(n, j), with  $j \in \{1, 2\}$  as appropriate.

Case 2: Merge. In this case,  $p_{i+1} = 2$ , so  $E_{e_{i+1}}^{i+1}$  is a hyp formula of both  $\Sigma_1$  and  $\Sigma_2$ . The respective lines are  $\lambda_1$ ,  $Y, E_{e_{i+1}}^{i+1} \succ_{\mathfrak{C}} E_{e_{i+1}}^{i+1}$ , and  $\lambda_2$ ,  $Z, E_{e_{i+1}}^{i+1} \succ_{\mathfrak{C}} E_{e_{i+1}}^{i+1}$ . Append to  $\Sigma^i$  the sequence  $\langle s_{r_1}(X) \dots, s_{p_1}(X) \rangle$ , where  $\langle s_{r_1}(Y) \dots, s_{p_1}(Y) \rangle$  is the top of  $\sigma_1(\lambda_1)$ . Then, append the sequence  $\langle s_{r_2}(X) \dots, s_{p_2}(X) \rangle$ , where  $\langle s_{r_2}(Z) \dots, s_{p_2}(Z) \rangle$  is the truncated top of  $\sigma_2(\lambda_2)$ . As in the previous case, let  $\lambda_j$  be the greatest line number from  $\Sigma_j$  such that  $f^i(\lambda_j, j)$  is defined. Set  $f^{i+1}(n, j) = f^i(n, j)$ , where  $n \leq \lambda_j$ . For each line  $s_q(Y)$  in the top of  $\sigma_1(\lambda_1)$ , set f(q, 1) = m, where m is the line number of the newly added  $s_q(X)$  in  $\Sigma^{i+1}$ . For each line  $s_q(Z)$  in the truncated top of  $\sigma_2(\lambda_2)$ , set f(q, 2) = m, where m is the line number of the newly added  $s_q(X)$  in  $\Sigma^{i+1}$ . Additionally, set  $f(\lambda_2 + 1, 2) = f(\lambda_1 + 1, 1)$ . Finally, each justification number n in  $\Sigma^{i+1}$  is replaced by f(n, j), with  $j \in \{1, 2\}$  as appropriate.

The preceding construction is the weave of  $\Sigma_1$  and  $\Sigma_2$ ,  $\mathcal{W}(\Sigma_1, \Sigma_2)$ , yielding  $\Sigma^{\ell}$  as the output.

There are two important features of  $W(\Sigma_1, \Sigma_2)$ . First, it is a cascade proof. Each stage produces a sequence of lines that satisfies the conditions on being a proof with appropriate rules and justifications. Second, since no indices are changed, if the proofs  $\Sigma_1, ..., \Sigma_n$  are compatible, then  $W(\Sigma_1, \Sigma_2), \Sigma_3, ..., \Sigma_n$  are compatible.

While the details of weaving can get a bit involved, the idea behind the construction is fairly straightforward. Given two cascade proofs, one orders their open assumptions so that the hyp indices are increasing. One proceeds down the sequence, pulling in steps, with modifications to the sequence of assumptions, from one or both proofs,

297

<sup>&</sup>lt;sup>21</sup> If  $f^{i}(\cdot, j)$  is undefined for every input, then set  $\lambda_{i} = 0$ .

ī.

Fig. 2. Weave case 1 in graphical form.

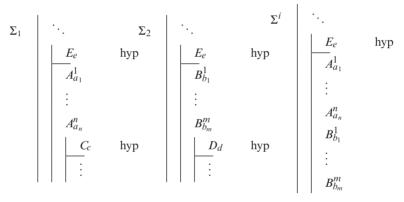


Fig. 3. Weave case 2 in graphical form.

according to the current member of the sequence of assumptions,  $(E_e, p)$  In the first case, one inserts the steps in the top of the subproof beginning with  $E_e$ , with the possible addition of more hypotheses. In the second, one inserts the steps in the top of the subproof beginning with  $E_e$  in one proof, and then the steps in the truncated top of the subproof starting with  $E_e$  in the other, "merging" the two hyp steps for  $E_e$  in the weave. Figure 2 presents the graphical form of case 1, with the leftmost proof being  $\Sigma^i$ , the middle displaying the relevant subproof top from  $\Sigma_j$ , with  $j \in \{1,2\}$ , and the rightmost presenting  $\Sigma^{i+1}$ . Figure 3 presents the graphical form of case 2, with the tops of the two displayed subproofs from  $\Sigma_1$  and  $\Sigma_2$  on the left merged to produce the subproof of  $\Sigma^i$  whose top is on the right.<sup>22</sup>

I can now proceed to the translation. One translates the tree proof  $\Pi$  from the leaves downwards to the conclusion. There are a several cases, depending on the rule used. It will be helpful to bear in mind that, since  $\Pi$  is a tree proof, the proof as a whole obeys the index discipline on assumptions.

<sup>&</sup>lt;sup>22</sup> The diagonal dots, ..., indicate that steps of the proofs may change in depth. Following diagonal dots, the number of vertical lines need not represent the depth of a line.

- Case: Assumption.  $\Pi$  is an assumption,  $A_{\{k\}}$ . This translates to a hyp step  $A_{\{k\}} \succ_{\mathfrak{C}} A_{\{k\}}$ .
- Case: One premise, no discharge. This case covers  $(\&E), (\lor I)$ , and (Dist). The case is the following.

$$\frac{\Pi}{B_a}(\rho)$$

By the inductive hypothesis, there is a translation  $\Sigma$  of  $\Pi$ , ending with the following line.

$$X \succ_{\mathfrak{C}} A_a$$

i

This is then extended as follows

$$i \qquad X \succ_{\mathfrak{C}} A_a$$
$$i+1 \qquad X \succ_{\mathfrak{C}} B_a \qquad \rho i$$

Case:  $(\rightarrow I)$ . The case is the following.

$$\frac{\overline{A_{\{k\}}}}{\Pi} \stackrel{(1)}{\underbrace{B_a}}{\underbrace{A \to B_{a-\{k\}}}} (\to I), (1)$$

By the inductive hypothesis, there is a translation  $\Sigma'$  of  $\Pi$ , ending with the following.

$$i \quad X, A_{\{k\}} \succ_{\mathfrak{C}} A_{\{k\}}$$
$$\vdots$$
$$j \quad X, A_{\{k\}} \succ_{\mathfrak{C}} B_a$$

Note that since k is the greatest integer among the indices on open assumptions,  $A_{\{k\}}$  heads the innermost subproof of  $\Sigma'$ . One can then extend  $\Sigma'$  with the following line.

$$j+1$$
  $X \succ_{\mathfrak{C}} A \to B_{a-\{k\}} \longrightarrow \mathbf{I} \ i-j$ 

Let this new proof be  $\Sigma$ , which is the desired translation. Case:  $(\rightarrow E)$ . The case is the following.

$$\frac{\Pi_1 \qquad \Pi_2}{\underline{A \to B_b \qquad A_a}}_{B_a \cup b} (\to E)$$

By the inductive hypothesis, there are translations  $\Sigma_1$  and  $\Sigma_2$  of  $\Pi_1$  and  $\Pi_2$ , respectively, ending with the following.

$$\begin{array}{ccc} \Sigma_1 & \Sigma_2 \\ j_1 & X \succ_{\mathfrak{C}} A \to B_b & j_2 & Y \succ_{\mathfrak{C}} A_a \end{array}$$

By the index discipline on the trees, the indices on open hypotheses of  $\Sigma_1$  and  $\Sigma_2$  satisfy the conditions on indices and they are compatible.<sup>23</sup> Construct the weave  $\mathcal{W}(\Sigma_1, \Sigma_2)$ . There are then two subcases, depending on whether  $\max(b) < \max(a)$ . In case  $\max(b) < \max(a)$ ,  $\mathcal{W}(\Sigma_1, \Sigma_2)$  end with the following, where U is the initial subsequence of  $\mu(X, Y)$  obtained at line *i*.

$$i \qquad U \succ_{\mathfrak{C}} A \to B_b$$

$$\vdots$$

$$n \qquad \mu(X, Y) \succ_{\mathfrak{C}} A_a$$

This is then extended with

$$n+1 \qquad \mu(X,Y) \succ_{\mathfrak{C}} A \to B_b \qquad \text{reit } i$$
  
$$n+2 \qquad \mu(X,Y) \succ_{\mathfrak{C}} B_{a \cup b} \qquad \to \mathbf{E} \ n, n+1$$

The step satisfies Brady's conditions, if any, because the step in the tree proof satisfied them. The subcase in which  $\max(b) \ge \max(a)$  is similar, with a different reit step.

Case: (&I). The case is the following.

$$\frac{\Pi_1 \qquad \Pi_2}{A_a \qquad B_a} (\&I)$$

By the inductive hypothesis, there are translations  $\Sigma_1$  and  $\Sigma_2$  of  $\Pi_1$  and  $\Pi_2$ , respectively, ending in the following.

$$\begin{array}{ccc} \Sigma_1 & \Sigma_2 \\ j_1 & X \succ_{\mathfrak{C}} A_a & j_2 & X \succ_{\mathfrak{C}} B_a \end{array}$$

Then,  $\mathcal{W}(\Sigma_1, \Sigma_2)$  ends with the following.

$$n \qquad X \succ_{\mathfrak{C}} A_a$$

$$\vdots$$

$$m \qquad X \succ_{\mathfrak{C}} B_a$$

 $\mathcal{W}(\Sigma_1, \Sigma_2)$  is then extended with the following.

The result is the desired translation.

Case:  $(\lor E)$ . The case is the following.

By the inductive hypothesis, there are translations  $\Sigma_1$ ,  $\Sigma_2$ , and  $\Sigma_3$  of  $\Pi_1$ ,  $\Pi_2$ , and  $\Pi_3$ , respectively, ending with the following.

<sup>&</sup>lt;sup>23</sup> The assumptions about the indices in trees ensure that translations of multipremise rules can be used in weaves. I will omit this in later steps with multipremise rules.

$$m+1 \quad X \succ_{\mathfrak{C}} A_a \qquad \text{reit } n$$

$$m+2 \quad X \succ_{\mathfrak{C}} A \& B_a \qquad \& \mathbf{I} \ m+1, m$$

$$\frac{\Pi_1 \qquad \Pi_2 \qquad \Pi_3}{A \rightarrow C_b \qquad B \rightarrow C_b \qquad A \lor B_a}_{C_a \cup b} (\lor E)$$

$$\Sigma_1 \qquad \Sigma_2 \qquad \Sigma_3$$

$$j_1 \qquad X \succ_{\mathfrak{C}} A \to C_b \qquad j_2 \qquad X \succ_{\mathfrak{C}} B \to C_b \qquad j_3 \qquad Y \succ_{\mathfrak{C}} A \lor B_a$$

 $\mathcal{W}(\Sigma_1, \Sigma_2)$  ends with the following.

$$n_1 \qquad X \succ_{\mathfrak{C}} A \to C_b$$

$$\vdots$$

$$n_2 \qquad X \succ_{\mathfrak{C}} B \to C_b$$

There are then two cases, depending on whether  $\max(b) < \max(a)$  or not. Only one,  $\max(b) < \max(a)$ , will be presented here, as they are similar.  $\mathcal{W}(\mathcal{W}(\Sigma_1, \Sigma_2), \Sigma_3)$  then ends with the following lines, where U is the initial subsequence of  $\mu(X, Y)$  obtained at lines  $m_1$  and  $m_2$ .

$$m_1 \qquad U \succ_{\mathfrak{C}} A \to C_b$$

$$\vdots$$

$$m_2 \qquad U \succ_{\mathfrak{C}} B \to C_b$$

$$\vdots$$

$$m \qquad \mu(X, Y) \succ_{\mathfrak{C}} A \lor B_a$$

This is extended by the following lines.

$$\begin{array}{ll} m+1 & \mu(X,Y)\succ_{\mathfrak{C}}A \to C_b & \text{reit } m_1 \\ \\ m+2 & \mu(X,Y)\succ_{\mathfrak{C}}B \to C_b & \text{reit } m_2 \\ \\ m+3 & \mu(X,Y)\succ_{\mathfrak{C}}C_{a\cup b} & \forall \mathbf{E}\ m,m+1,m+2 \end{array}$$

The result is the desired translation. The translation satisfies Brady's conditions, if any, because the original tree proof satisfied them.Call the translation just described  $\mathbb{C}$ . So  $\mathbb{C}(\Pi) = \Sigma$ . We have the following theorem.

THEOREM 2. Let  $\Pi$  be a tree proof of  $A_{a_1}^1, \ldots, A_{a_n}^n \vdash_{\mathfrak{T}} B_b$ . Then  $\mathbb{C}(\Pi) = \Sigma$  is a cascade proof of  $A_{a_1}^1, \ldots, A_{a_n}^n \vdash_{\mathfrak{C}} B_b$ .

Tree proofs translate to cascade proofs, and cascade proofs translate to tree proofs. Brady (1984) shows that the cascade systems are sound and complete for the relevant logic X, where X is a logic listed in 1. As Brady's results carry over to the cascade systems of this paper, it follows that the tree system for the relevant logic X are sound and complete for X.

**§5.** Conclusions. In this paper, I have provided indexed tree natural deduction proof systems for a range of relevant logics with cascade proof systems. The translations above demonstrate the adequacy of these systems and, I hope, shed some light on the mechanisms of the index discipline and restrictions in the Anderson–Belnap–Brady systems. I will close with some comments on the work so far, beginning with a discussion of the trade offs of the two kinds of natural deduction systems, followed by accommodating negation and the logic E, and ending with a discussion of dependency as found in the indexed systems, contrasted with Lemmon-style systems.

There is a trade off in the move from tree to cascade or conversely: In the cascade systems, the dependencies of the assumptions are made clear by the subproof structure, whose linearity requires abandoning immediate connection between premises and conclusions of rules; in contrast, the tree systems enforce an immediate connection between premises and conclusions but obscure the dependencies of assumptions by spreading them throughout the tree structure.<sup>24</sup> Given these trade offs, one might wonder what advantages are to be gained from moving to a tree proof system. There are two potential payoffs that I would like to highlight. First, fusion is a very natural connective in the context of trees whose inferential role is, perhaps, less clear in cascade systems. The rules for fusion are the following.<sup>25</sup>

$$\frac{A_a}{A \circ B_a \cup b} (\circ I) \qquad \qquad \frac{A \to (B \to C)_{\emptyset} \qquad A \circ B_a}{C_a} (\circ E)$$

For logics weaker than R, conditions will need to be placed on the indices in both rules, which I leave open here. In the context of tree proofs, the contribution of fusion may be easier to see than in the context of cascade proofs.

Second, techniques for proving normalization theorems are well known for tree natural deduction systems.<sup>26</sup> Normalization for cascade systems, while not unknown, is not as common.<sup>27</sup> These systems may facilitate normalization proofs for some of these relevant logics, particularly if (*Dist*) is dropped. Dropping (*Dist*) results in the distribution-less logic corresponding to dropping the distribution axiom from the Hilbert-style system or the distribution rule from the cascade system.

Next, a comment on extensions of the systems. In 1, I indicated that negation rules can be added without any problems. The translations extend to negation rules in a

<sup>&</sup>lt;sup>24</sup> For more on the trade offs involved in cascade and tree systems, see Restall (2014) and Hazen & Pelletier (2014).

<sup>&</sup>lt;sup>25</sup> I have presented the  $(\circ E)$  rule that I did, rather than a general elimination-style rule to keep the correspondence with cascade systems. See Read (2010, 2014) or Jacinto & Read (2017) for more on general elimination-style rules.

<sup>&</sup>lt;sup>26</sup> See, for example, Prawitz (1965), Raggio (1965), Medeiros (2006), and Francez (2014), among others.

<sup>&</sup>lt;sup>27</sup> See Brady (2006, 2010) for detailed examples. See Fitch (1952, 115ff.) for an early example.

straightforward way. Thus, a large swath of the propositional logics studied by Brady (1984) can be given tree systems.<sup>28</sup>

Returning to the motivations laid out in the introduction, one might wonder about E, which is conspicuous in its absence from the bulk of this paper. Anderson and Belnap's cascade system for E is obtained by imposing a restriction on the rule of reiteration, which captures the modal aspect of E's conditional. Brady (1984) does not offer a cascade system for E, since those systems are obtained by restrictions on the rules  $\rightarrow$ E and  $\lor$ E rather than on reit. There are two ways one can proceed for giving a tree natural deduction system for E. The first is to use one of the systems discussed by Standefer & Brady (2019) as the basis for a tree system. The second is to use the tree system of Standefer (2018). These offer different views of natural deduction for E.

Finally, I will close with a comment about the indices, which play such an important role in Anderson–Belnap–Brady systems, as well as the tree systems presented here. Some logicians, such as Prawitz, give systems that do not use indices, opting for side conditions on rules. The side conditions can, sometimes, be complex and nonlocal. The indices have the advantage of being local annotations, in the sense of going along with each formula, and permitting simple rules. To check whether a rule can be applied, one checks the indices on the premises of the rule, rather than checking a global condition on the proofs. The indices are simple, in the sense that they are just sets of natural numbers. Other approaches to relevant logic tend to use more general structures to indicate dependencies on premises.<sup>29</sup> The indices are, further, not model-theoretically loaded. They do not need to be understood as representing features of a model, as usually happens in indexed tableau systems.<sup>30</sup> This is not to say that the indices must be treated purely instrumentally and without philosophical import.<sup>31</sup> The indices here track use and dependence, important concepts in relevant logic.

The main advantage of the indices over the Lemmon-style systems of, say, Slaney (1990) is that the structure of dependencies is simpler. A set of natural numbers has a natural representation as the sequence of those numbers in increasing order, as opposed to the more general groupoid structures used by Slaney. Assumptions corresponding to later numbers depend on the assumptions corresponding to earlier numbers; newer assumptions come later. With that in mind, we can revisit Brady's conditions from 1.

The rule  $\rightarrow$ E will be the rule of interest below, with  $A \rightarrow B_b$  the major premise and  $A_a$  the minor. The condition for B requires that the minor premise depend on only one assumption, which is newer than anything the major premise depends on, so the dependencies of the conclusion extend those of the major premise by one. The condition added for DJ permits the dependencies of the major and minor premises to overlap significantly, requiring that the minor premise depends on what the major premise does, with the addition of one newer assumption. The dependencies of the conclusion are then those of the minor premise. With both B and DJ, the dependencies

<sup>&</sup>lt;sup>28</sup> The question of whether these techniques can be extended to the quantifiers is an interesting one that I leave open. An additional question is how to accommodate the disjunctive rules considered by Brady.

<sup>&</sup>lt;sup>29</sup> See, for example, Slaney (1990), Read (1988), Restall (2000), or Mares (2017).

<sup>&</sup>lt;sup>30</sup> For examples, see Priest (2008), especially the chapter on relevant logic, or Mints (1992), for examples from modal logic. See Read (2015) for discussion of labels in proof systems from an inferentialist point of view.

<sup>&</sup>lt;sup>31</sup> See Mares (2004) for an interpretation of the indices in terms of situated inference. The indices can, in some cases, give rise to models, as is the case in Urquhart's semantics.

of the conclusion extend those of the major premise by one. The condition for T says that the minor premise must depend on newer assumptions than the major premise. The dependencies of the conclusion are interlaced in the manner of the weave construction.<sup>32</sup> The condition for RW requires that the dependencies be, in a sense, disjoint, and the conclusion gathers together these dependencies. Finally, R eschews additional restrictions, relying only on the strictures of  $\rightarrow$ I.

For the indexed natural deduction systems of this paper, there is a sense in which it is most natural to start with R, obtaining weaker logics by adding conditions on the dependencies being appropriate for  $\rightarrow E$  and  $\lor E$ . In contrast, the Lemmon-style systems "start" with the weaker logics, around B. They have rules operating on premise structures, which are not part of the cascade or tree systems of this paper. Stronger logics are obtained through the adoption of additional rules dealing with premise structure. These systems appear to provide a different view of dependence than that of the indexed systems, although a fuller discussion is beyond the scope of this paper.

**Acknowledgments** I would like to thank Greg Restall, Lloyd Humberstone, Ross Brady, and the audience at the Melbourne Logic Seminar for comments and discussion that have improved this paper. Thanks also to an anonymous referee at this journal for constructive feedback that greatly improved this paper. This research was supported by the Australian Research Council, Discovery Grant DP150103801.

#### BIBLIOGRAPHY

- Anderson, A. R., & Belnap, N. D. (1975). Entailment: The Logic of Relevance and Necessity, Vol. I. Princeton, NJ: Princeton University Press.
- Anderson, A. R., Belnap, N. D., & Dunn, J. M. (1992). Entailment: The Logic of Relevance and Necessity, Vol. 2. Princeton, NJ: Princeton University Press.
- Bimbó, K. (2006). Relevance logics. In Jacquette, D., editor. *Philosophy of Logic*, *Volume 5 of Handbook of the Philosophy of Science*, Amsterdam, Netherlands: Elsevier, pp. 723–789.
- Borghuis, T. (1998). Modal pure type systems. *Journal of Logic, Language and Information*, 7(3), 265–296.
- Brady, R. T. (1984). Natural deduction systems for some quantified relevant logics. *Logique Et Analyse*, **27**(8), 355–377.
- \_\_\_\_\_. (2006). Normalized natural deduction systems for some relevant logics I: The logic DW. *Journal of Symbolic Logic*, **71**(1), 35–66.
  - . (2010). Free semantics. *Journal of Philosophical Logic*, **39**(5), 511–529.
- Charlwood, G. (1981). An axiomatic version of positive semilattice relevance logic. *Journal of Symbolic Logic*, **46**(2), 233–239.
- Church, A. (1953). The weak theory of implication. *Journal of Symbolic Logic*, **18**(4), 326–326.
- Došen, K. (1992). The first axiomatization of relevant logic. Journal of Philosophical Logic, 21(4), 339–356.

<sup>&</sup>lt;sup>32</sup> Cf. the generalized transitivity axiom of Anderson & Belnap (1975, pp. 26–27).

- Dunn, J. M., & Restall, G. (2002). Relevance logic. In Gabbay, D. M. & Guenthner, F., editors. *Handbook of Philosophical Logic* (Second edition), Vol. 6, Dordrecht: Springer, pp. 1–136.
- Fitch, F. B. (1952). Symbolic Logic. New York, NY: Ronald Press Co.

Francez, N. (2014). Bilateral relevant logic. Review of Symbolic Logic, 7(2), 250–272.

- Giambrone, S., & Urquhart, A. (1987). Proof theories for semilattice logics. *Zeitschrift fur mathematische Logik und Grundlagen der Mathematik*, **33**(5), 433–439.
- Hazen, A. P., & Pelletier, F. J. (2014). Gentzen and Jaśkowski natural deduction: Fundamentally similar but importantly different. *Studia Logica*, **102**(6), 1103–1142.
- Humberstone, L. (1988). Operational semantics for positive R. *Notre Dame Journal of Formal Logic*, **29**, 61–80.
  - \_\_\_. (2011). The Connectives. Cambridge, MA: MIT Press.
- Indrzejczak, A. (2010). *Natural Deduction, Hybrid Systems and Modal Logics*. New York, NY: Springer.
- Jacinto, B., & Read, S. (2017). General-elimination stability. *Studia Logica*, **105**, 361–405
- Jaśkowski, S. (1934). On the rules of suppositions in formal logic. *Studia Logica*, **1**, 5–32. Reprinted in *Polish Logic 1920–1939*, edited by S. McCall, Oxford University Press, pp. 232–258.
- Leivant, D. (1979). Assumption classes in natural deduction. Zeitschrift fur mathematische Logik und Grundlagen der Mathematik, **25**(1-2), 1–4.
- Lemmon, E. J. (1965). Beginning Logic. New Zealand: Nelson.
- Mares, E. D. (2017). An informational interpretation of weak relevant logic and relevant property theory. *Synthese*, 1–23. Forthcoming.

\_\_\_\_\_. (2004). Relevant Logic: *A Philosophical Interpretation*. Cambridge, MA: Cambridge University Press.

- Medeiros, M. D. P. N. (2006). A new S4 classical modal logic in natural deduction. *Journal of Symbolic Logic*, **71**(3), 799–809.
- Meyer, R. K., Martin, E. P., Giambrone, S., & Urquhart, A. (1988). Further results on proof theories for semilattice logics. *Mathematical Logic Quarterly*, **34**(4), 301–304.
- Mints, G. (1992). A Short Introduction to Modal Logic. Stanford, CA: CSLI Publications.
- Moh, S.-K. (1950). The deduction theorems and two new logical systems. *Methodos*, **2**, 56–75.
- Pelletier, F. J. (1999). A brief history of natural deduction. *History and Philosophy of Logic*, **20**(1), 1–31.
- Prawitz, D. (1965). *Natural Deduction: A Proof-Theoretical Study*. Stockholm: Almqvist and Wicksell.
- Priest, G. (2008). An Introduction to Non-Classical Logic: From If to Is. Cambridge, MA: Cambridge University Press.
- Raggio, A. (1965). Gentzen's Hauptsatz for the systems NI and NK. *Logique Et Analyse*, **8**, 91–100.
- Read, S. (1988). *Relevant Logic: A Philosophical Examination of Inference*. Oxford, UK: Basil Blackwell.
- . (2010). General-elimination harmony and the meaning of the logical constants. *Journal of Philosophical Logic*, **39**(5), 557–576.
- . (2014). General-elimination harmony and higher-level rules. In Wansing, H., editor. *Dag Prawitz on Proofs and Meaning, Outstanding Contributions to Logic*, pp. 293–312. New York, NY: Springer.

. (2015). Semantic pollution and syntactic purity. *Review of Symbolic Logic*, **8**(4), 649–661.

- Routley, R., Plumwood, V., Meyer, R. K., & Brady, R. T. (1982). *Relevant Logics and Their Rivals*, Vol. 1. Chanhassen, MN: Ridgeview.

Slaney, J. (1990). A general logic. Australasian Journal of Philosophy, 68(1), 74-88.

Standefer, S. (2018). Trees for E. Logic Journal of the IGPL, 26(3), 300–315.

\_\_\_\_\_. (2019). Translations between Gentzen–Prawitz and Jaśkowski–Fitch natural deduction proofs. *Studia Logica*, **107**, 1103–1134.

Standefer, S., & Brady, R. T. (2019). Natural deduction systems for E. Logique Et Analyse, 61(242), 163–182.

Tennant, N. (2012). Cut for core logic. *Review of Symbolic Logic*, 5(3), 450–479.
. (2015). The relevance of premises to conclusions of core proofs. *Review of Symbolic Logic*, 8(4), 743–784.

. (2017). Core Logic. Oxford, UK: Oxford University Press.

Urquhart, A. (1972). Semantics for relevant logics. *Journal of Symbolic Logic*, **37**(1), 159–169.

von Plato, J. (2017). From Gentzen to Jaskowski and back: Algorithmic translation of derivations between the two main systems of natural deduction. *Bulletin of the Section of Logic*, 46(1/2), 65–73.

SCHOOL OF HISTORICAL AND PHILOSOPHICAL STUDIES THE UNIVERSITY OF MELBOURNE PARKVILLE, VIC 3010, AUSTRALIA

*E-mail*: standefer@gmail.com *URL*: http://www.standefer.net