Converting $\mathcal{ALC}$ Connection Proofs into $\mathcal{ALC}$ Sequents

Eunice Palmeira  
Federal Institute of Alagoas  
Maceió - AL, Brazil  
eunicepalmeira@ifal.edu.br

Fred Freitas  
Federal University of Pernambuco  
Recife - PE, Brazil  
fred@cin.ufpe.br

Jens Otten  
University of Oslo  
Oslo, Norway  
jeotten@ifi.uio.no

The connection method has earned good reputation in the area of automated theorem proving, due to its simplicity, efficiency and rational use of memory. This method has been applied recently in automatic provers that reason over ontologies written in the description logic $\mathcal{ALC}$. However, proofs generated by connection calculi are difficult to understand. Proof readability is largely lost by the transformations to disjunctive normal form applied over the formulae to be proven. Such a proof model, albeit efficient, prevents inference systems based on it from effectively providing justifications and/or descriptions of the steps used in inferences. To address this problem, in this paper we propose a method for converting matricial proofs generated by the $\mathcal{ALC}$ connection method to $\mathcal{ALC}$ sequent proofs, which are much easier to understand, and whose translation to natural language is more straightforward. We also describe a calculus that accepts the input formula in a non-clausal $\mathcal{ALC}$ format, what simplifies the translation.

1 Introduction

Description Logics (DLs) \cite{hobbs1995description} are a family of knowledge representation formalisms considered as a fundamental foundation for the Semantic Web, as it constitutes the formalism underlying the Web Ontology Language (OWL) language. DL is an expressive, decidable subset of First Order Logic (FOL), successfully applied in several areas. DL provides a precise and unambiguous meaning to DL descriptions due to its formal semantics, and fast reasoners have been produced to the many fragments available \cite{conte2007using}.

One of them, the $\mathcal{ALC}$ $\theta$-Connections Calculus, and its automated reasoner RACCOON (Reasoner based on the Connection Calculus Over ONtologies), is based on the Connection Method \cite{feldman1992connection,ottens1999connection}, and was specifically developed to infer over the Description Logic $\mathcal{ALC}$ \cite{feldman1992connection,ottens1999connection}. The calculus includes typical DL features and techniques, such as notation without variables, absence of Skolem functions/unification and, inclusion of a blocking rule to handle cycles, which guarantees termination to make for the case of cyclic ontologies. The Connection Calculus has earned good reputation in the area of automated theorem proving due to its simplicity, efficiency and rational use of memory. The method represents formulae as matrices, whose columns are conjunctive clauses; its proof procedure consists of horizontally traversing paths through the matrix in order to connect complimentary literals (e.g., $L$ with its complement $\neg L$). A pair $\{L, \neg L\}$, is called a connection, which corresponds to the validity the path being checked. Thus, a formula is valid if every path through the matrix corresponding to it has a connection.

Both calculi mentioned above, before attempting to find a proof, convert a formula into a disjunctive normal form. The translation to this clausal form often obscures the structure of the original formula and transforms some simple theorem proofs into difficult ones\cite{hobbs1995description}. In complex cases, the deductions’ premise(s) and conclusion can no longer be clearly identified, once the transformation has been applied \cite{hobbs1995description}. Thus, proof readability and understandability is largely lost, and consequently, it becomes quite difficult to provide justifications and/or descriptions of the steps used during inferences.

The $\theta$-Non-clausal $\mathcal{ALC}$ $\theta$-Connection Calculus is based on the $\mathcal{ALC}$ $\theta$-Connection Calculus and works directly on the structure of the original formula, thus avoiding the translation into a clausal form.
Nevertheless, its proof format is still not intuitive, once, like other connection calculi, it consists of a set of complementary pairs found in each path through the matrix, when the formula is valid.

The motivation of this work is to make a connection proof for ALC more readable so that, in a near future, justifications can be generated automatically in natural language. Therefore, this article proposes a conversion method that translates non-clausal ALC \( \theta \)-connection proofs into ALC sequent proofs. Sequent calculi have a more friendly proof representation than connection calculus; it conveys proofs in a formal logic argument style, where each proof line is a conditional tautology. Such translation should therefore contribute to a better user interaction with DL reasoners based on the Connection Method.

The DL ALC is presented in the next section; Section 3 brings an ALC non-clausal Connection Calculus for ALC; Section 4 introduces the ALC Sequent Calculus, to which proofs will be translated; the conversion process and its main concepts in Section 5; an overview of the main algorithms for the conversion method with its computational complexities in Section 6; and conclusions in Section 7.

2 The Description Logic ALC

An ontology \( O \) in ALC is a set of axioms over a signature \((N_C, N_R, N_O)\), where \( N_C \) is the set of concept names (unary predicate symbols), \( N_R \) is the set of role or property names (binary predicate symbols); \( N_O \) is the set of individual names (constants) \([1]\). Concept expressions are inductively defined as follows. \( N_C \) includes \( \top \), the universal concept that subsumes all concepts, and \( \bot \), the bottom concept subsumed by all concept names belong to \( N_C \). If \( r \in N_R \) is a role and \( C, D \in N_C \) are concepts, then the following formulæ are also concepts: (i) \( C \sqcap D \), (ii) \( C \sqcup D \), (iii) \( \neg C \), (iv) \( \forall r.C \); (v) \( \exists r.C \).

A knowledge base in DL consists of a set of basic axioms (TBox), and a set of axioms specific to a particular situation (ABox). Two axiom types are allowed in a TBox \( T \): (i) \( C \sqsubseteq D \); (ii) \( C \sqsubseteq D \), standing for \( C \sqsubseteq D \) and \( D \sqsubseteq C \). An ABox \( A \) w.r.t. a TBox \( T \) is a finite set of assertions of two types: (i) a concept assertion is a statement of the form \( C(a) \), where \( a \in N_O \), \( C \in N_C \) and (ii) a role assertion \( r(a,b) \), where \( a,b \in N_O \), \( r \in N_R \). An ALC formula is either an axiom or an assertion; an ontology \( O \) is an ordered pair \((T,A)\). The semantics of concepts and ontologies is defined in the usual way - see, e.g., \([1]\).

3 The Non-clausal ALC \( \theta \)-Connection Calculus

**Definition 1.** *(Query).* A query \( O \models \alpha \) is an ALC formula to be proven valid, where \( O \) is an ALC ontology, and \( \alpha \) is either a TBox or an ABox axiom to be proven a logical consequence from \( O \).

**Definition 2.** *(Literal, clause, matrix).* ALC Literals are atomic concepts or roles, possibly negated or instantiated in the form \( L \) or \( \neg L \). An ALC disjunction is either a literal \( L \), a disjunction \((E_0 \sqcup E_1)\) or an universal restriction \( \forall r.E_0 \). An ALC conjunction is either a literal \( L \), a conjunction \((E_0 \sqcap E_1)\) or an existential restriction \( \exists r.E_0 \), where \( E_0 \) and \( E_1 \) are expressions of arbitrary concepts (see DLs and its Mapping to FOL in \([2]\)). Clauses are conjunctions of literals and matrices in the form \( L_1 \sqcap \ldots \sqcap L_m \), where each \( L_i \) is a literal or a matrix. A matrix of a formula (in DNF) is its representation as a set \( \{C_1, \ldots, C_n\} \), where each \( C_i \) is a clause.

**Definition 3.** *(Formula with polarity).* A formula with polarity, denoted by \( F^p \), consists of a formula \( F \) and a polarity \( p \), where \( p \in \{0,1\} \), that is, 0 is positive and 1 is negative. This concept is used to denote negation in a matrix, i.e. literals or matrices \( A \) and \( \neg A \) are represented by \( A^0 \) and \( A^1 \), respectively.

**Definition 4.** *(ALC Non-Clausal Matrix).* An ALC non-clausal matrix is a set of clauses in which a clause is a set of literals and matrices. Let \( F \) be a formula and \( p \) be a polarity. The matrix of \( F^p \), denoted
by \( M(F^p) \), is inductively defined according to Table 1 which indicates how the polarity is inherited by the (sub-)matrices of an \( F^p \). The matrix of \( F^p \) is the matrix \( M(F^0) \). Literals or (sub-)matrices involved in a universal restriction (\( \forall r.C \)) or in an existential restriction (\( \exists r.C \)) are underlined in the matrix.

### Table 1: Matrix of an \( \mathcal{ALC} \) formula \( F^p \).

<table>
<thead>
<tr>
<th>Type</th>
<th>( F^p )</th>
<th>( M(F^p) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atomic</td>
<td>( A^0 )</td>
<td>{{}}</td>
</tr>
<tr>
<td></td>
<td>( A^1 )</td>
<td>{{A}}</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>( \neg C^0 )</td>
<td>{{C}}</td>
</tr>
<tr>
<td></td>
<td>( \neg C^1 )</td>
<td>{{C}}</td>
</tr>
<tr>
<td></td>
<td>( (\lor D)^1 )</td>
<td>{{{}}{{}}}</td>
</tr>
<tr>
<td></td>
<td>( (\land D)^0 )</td>
<td>{{{}}{{}}}</td>
</tr>
<tr>
<td></td>
<td>( (\land D)^0 )</td>
<td>{{{}}{{}}}</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type</th>
<th>( F^p )</th>
<th>( M(F^p) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta )</td>
<td>( C \lor D^0 )</td>
<td>{{{}}{{}}}</td>
</tr>
<tr>
<td></td>
<td>( C \land D )</td>
<td>{{{}}{{}}}</td>
</tr>
<tr>
<td>( \gamma )</td>
<td>( (\forall r.D)^1 )</td>
<td>{{{}}{{}}}</td>
</tr>
<tr>
<td>( \delta )</td>
<td>( (\exists r.D)^0 )</td>
<td>{{{}}{{}}}</td>
</tr>
</tbody>
</table>

**Definition 5.** (Positive) Graphical Representation of the Matrix. In the (positive) graphical representation of a matrix, its clauses are arranged horizontally, while the literals and (sub-)matrices of each clause are arranged vertically. The restrictions are represented by solid lines; when a restriction involves more than one clause, its literals are indexed in the bottom with the same index in the matrix column in the written representation, for example, the notation \( L_1 \) (see example 1): restrictions with indexes are represented with horizontal lines; restrictions without indexes with vertical lines.

**Example 1.** (Query, clause, \( \mathcal{ALC} \) non-clausal matrix, formula with polarity, graphical representation of a matrix). The query \( F_1 = \exists \text{pet}.\text{Cat} \subseteq \text{CatOwner}, \text{OldLady} \subseteq \exists \text{hasPet}.\text{Animal} \cap \forall \text{hasPet}.\text{Cat} \models \text{OldLady} \subseteq \text{CatOwner} \) is read in FOL as:

\[
\forall x(\exists y \text{hasPet}(x,y) \land \text{Cat}(y)) \rightarrow \text{CatOwner}(x) \\
\forall z(\text{OldLady}(z) \rightarrow \exists v(\exists \text{hasPet}(z,v) \land \text{Animal}(v))) \\
\wedge \forall k(\exists \text{hasPet}(z,k) \rightarrow \text{Cat}(k))
\]

and is represented by the FOL matrix (\( a \) is a Skolem term, \( f \) a function symbol):

\[
\{\{\text{hasPet}^0, \text{Cat}^0, \text{CatOwner}^1\}, \{\text{OldLady}^0, \{\text{hasPet}^1\}, \{\text{Animal}^1\}, \{\text{hasPet}^0, \text{Cat}^1\}\}\}, \{\text{OldLady}^0, \{\text{Cat}^1\}\}
\]

and by the following \( \mathcal{ALC} \) non-clausal matrix \( M_1 \), which is defined according to Table 1 (column indices relate the two clauses involved in a same restriction; variables are omitted as they are specified implicitly):

\[
\{\{\text{hasPet}^0, \text{Cat}^0, \text{CatOwner}^1\}, \{\text{OldLady}^0, \{\text{hasPet}^1\}, \{\text{Animal}^1\}, \{\text{hasPet}^0, \text{Cat}^1\}\}\}, \{\text{OldLady}^0, \{\text{Cat}^1\}\}
\]

So, its graphical representation of the \( M_1 \) is:

\[
\begin{pmatrix}
\text{hasPet}^0 & \text{Cat}^0 & \text{CatOwner}^1 \\
\text{OldLady}^0 \\
\end{pmatrix}
\begin{pmatrix}
\text{hasPet}^1 \\
\text{Animal}^1 \\
\end{pmatrix}
\begin{pmatrix}
\text{hasPet}^0 & \text{Cat}^1 \\
\text{OldLady}^0 \\
\end{pmatrix}
\begin{pmatrix}
\text{CatOwner}^0 \\
\end{pmatrix}
\]

Matrices of the form \( M = \{ \ldots \{ C_1, \ldots, C_n \}, \ldots \} \) can be simplified to \( M' = \{ \ldots, C_1, \ldots, C_n, \ldots \} \), where \( C_1, \ldots, C_n \) are clauses. Clauses of the form \( C = \{ \ldots, \{ M_1, \ldots, M_m \}, \ldots \} \) can be simplified to \( C' = \{ \ldots, M_1, \ldots, M_m, \ldots \} \), where \( M_1, \ldots, M_m \) are matrices.
Definition 6. (Path). A path through a matrix \( M = \{C_1, \ldots, C_n\} \) is a set of literals containing a literal \( L_i \) of each clause \( C_i \in M \), i.e., \( \bigcup_{i=1}^{n} \{L_i\} \) with \( L_i \in C_i \). A path through a matrix \( M \) (or a clause \( C \)) is inductively defined as follows. The (only) path through a literal \( L \) is \( \{L\} \). If \( p_1, \ldots, p_n \) are paths through the clauses \( C_1, \ldots, C_n \), respectively, then \( p_1 \cup \ldots \cup p_n \) is a path through the matrix \( M = \{C_1, \ldots, C_n\} \). If \( p_1, \ldots, p_n \) are paths through the matrices/literals \( M_1, \ldots, M_n \), respectively, then \( p_1, \ldots, p_n \) are also paths through the clause \( C = \{M_1, \ldots, M_n\} \).

Definition 7. (Connection, \( \theta \)-substitution, \( \theta \)-complementary connection). A connection is a pair of literals \( \{E, \neg E\} \) with the same concept/role name, but different polarities. A \( \theta \)-substitution assigns to each (possibly omitted) variable an individual or another variable (in the whole matrix). A \( \theta \)-complementary connection is a pair of \( \mathcal{ALC} \) literals \( \{E(x), \neg E(y)\} \) or \( \{p(x, v), \neg p(y, u)\} \), with \( \theta(x) = \theta(y), \theta(v) = \theta(u) \). The complement \( \bar{L} \) of a literal \( L \) is \( E \) if \( L = \bar{E} \). If \( p_1, \ldots, p_n \) are paths through the clauses \( C_1, \ldots, C_n \), then \( p_1 \cup \ldots \cup p_n \) is a path through the matrix \( M = \{C_1, \ldots, C_n\} \). If \( p_1, \ldots, p_n \) are paths through the matrices/literals \( M_1, \ldots, M_n \), respectively, then \( p_1, \ldots, p_n \) are also paths through the clause \( C = \{M_1, \ldots, M_n\} \).

Simple term unification without Skolem functions is used to calculate \( \theta \)-substitutions. The application of a \( \theta \)-substitution to a literal is an application to its variables, i.e. \( \theta(E) = E(\theta(x)) \) and \( \theta(r) = r(\theta(x), y) \), where \( E \) is an atomic concept and \( r \) is a role. Furthermore, \( x^\theta = \theta(x) \).

Example 2. (Path, Connection, \( \theta \)-substitution, \( \theta \)-complementary connection). In the matrix \( M_1 \) of Example 1, \( \{\text{hasPet}^1, \text{Animal}^1, \text{hasPet}^0, \text{OldLady}^1, \text{CatOwner}(a)^0\} \) and \( \{\text{Cat}, \text{hasPet}^1, \text{Animal}^1, \text{Cat}^1, \text{OldLady}^1, \text{CatOwner}(a)^0\} \) are some paths through \( M_1 \). \( \{\text{Cat}^0, \text{Cat}^1\} \) is a connection. \( \theta(\text{OldLady}^0) = \text{OldLady}(\theta(y))^0 \) and \( \theta(\text{hasPet}^0) = \text{hasPet}(\theta(y), x)^0 \), where \( \theta(y) = a \), are examples of \( \theta \)-substitution, and \( \{\text{OldLady}^0, \text{OldLady}(a)^1\} \) is a \( \theta \)-complementary connection.

Definition 8. (Set of concepts, Skolem condition). The set of concepts \( \tau(x) \) of a variable or individual \( x \) contains all concepts that were substituted/instantiated by \( x \) so far, i.e. \( \tau(x) \stackrel{\text{def}}{=} \{E(x) \in \text{Path}\} \), where \( E \) is a concept and \( E(x) \) is a substituted/instantiated literal coming from this concept. The Skolem condition ensures that at most one concept is underlined in the graphical matrix. The condition is formally stated as, \( \forall a \exists \{x^a \in \text{Path}\} \leq 1 \), with \( a \) a variable/individual, and \( i \) a column index.

Definition 9. (\( \alpha \)-Related Clause). Let \( C \) be a clause in a matrix \( M \) and \( L \) be a literal in \( M \). \( C \) is \( \alpha \)-related to \( L \), iff \( M \) contains (or is equal to) a matrix \( \{C_1, \ldots, C_n\} \) such that \( C = C_i \) or \( C_i \) contains \( C \), and \( C_j \) contains \( L \) for some \( 1 \leq i, j \leq n \) with \( i \neq j \). \( C \) is \( \alpha \)-related to all literals \( L \) in \( \mathcal{L} \).

Example 3. (\( \alpha \)-Related Clause) In the matrix of Example 1, \( \{\text{Animal}^1\} \) is \( \alpha \)-related to \( \{\text{hasPet}^0, \text{Cat}^1\} \).

Definition 10. (Parent Clause). Let \( M \) be a matrix and \( C \) be a clause in \( M \). The clause \( C' = \{M_1, \ldots, M_n\} \) in \( M \) is called the parent clause of \( C \) iff \( C \in M_i \), for some \( 1 \leq i \leq n \).

Example 4. (Parent Clause). In Example 1, \( \{\text{OldLady}^0, \{\text{hasPet}^1\}, \{\text{Animal}^1\}, \{\text{hasPet}^0, \text{Cat}^1\}\} \) is parent clause of \( \{\text{hasPet}^1\} \).

Definition 11. (Extension Clause). Let \( M \) be a matrix and \( P \) a path (be a set of literals). Then the clause \( C \) in \( M \) is an extension clause of \( M \) with respect to \( P \), iff either \( C \) contains a literal of \( P \), or \( C \) is \( \alpha \)-related to all literals of \( P \) occurring in \( M \) and if \( C \) has a parent clause, it contains a literal of \( P \).

In the extension rule of the \( \mathcal{ALC} \) \( \theta \)-Connection Calculus \( \{\text{S}, \text{I}\} \) the new subgoal clause (set of literals that need to be connected) is \( C_2 \setminus \{L_2\} \). In the non-clausal connection calculus the extension clause \( C_2 \) might contain clauses that are \( \alpha \)-related to \( L_2 \) and do not need to be considered for the new subgoal clause. Hence, these clauses can be deleted from the subgoal clause. The resulting clause is called the \( \beta \)-clause of \( C_2 \) with respect to \( L_2 \).
Definition 12. (β-Clause). Let $C = \{M_1, \ldots, M_n\}$ be a clause and $L$ be a literal in $C$. The β-Clause of $C$ with respect to $L$, denoted by $\beta_{\text{Clause}}(C)$, is inductively defined:

$$\beta_{\text{Clause}}(C) := \begin{cases} C \setminus \{L\} & \text{if } L \in C, \\ M_1, \ldots, M_{i-1}, \{C^\theta\}, M_{i+1}, \ldots, M_n & \text{otherwise}, \end{cases}$$

where $C' \in M_i$ contains $L$ and $C^\theta := \beta_{\text{Clause}}(C')$.

Example 5. (Extension Clause, β-Clause). In Example $7$ $C = \{\text{OldLady}^0, \{\text{hasPet}^1\}, \{\text{Animal}^1\}, \{\text{hasPet}^0\}, \text{Cat}^1\}$ is an extension clause with respect to $p = \{\text{CatOwner}(a)^0, \text{Cat}^0\}$, while the clause $\{\text{OldLady}^0, \{\text{hasPet}^1\}, \{\text{Animal}^1\}, \{\text{hasPet}^0\}\}$ is a β-Clause of $C$ with respect to $L = \text{Cat}^1$.

3.1 The Formal Non-Clausal ALC θ-Connection Calculus

Suppose we wish to entail if $O \models \alpha$ is valid using a direct method, like the Connection Method (CM). By the Deduction Theorem [3], we must then prove directly if $O \rightarrow \alpha$, or, in other words, if $\neg O \lor \alpha$ is valid. This opposes to classical refutation methods, like tableaux and resolution, which builds a proof by testing whether $O \cup \{\neg \alpha\} \models \bot$. Hence, in the CM, the whole knowledge base KB should be negated. Given $O = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, $\alpha_i$ being literal conjunctions in the clausal connection method, all (negated KB) formulae are converted to the Disjunctive Normal Form (DNF). A query then is the matrix $\theta$ valid. This opposes to classical refutation methods, like tableaux and resolution, which builds a proof from the $\alpha_1$ clauses only with literals, they can also contain matrices, and no conversion is needed. If every path contains a (θ-complementary) connection (representing a subformula $A \sqcup \neg A$ in a disjunction, what makes this disjunction valid), then the matrix is valid.

Definition 13. (Non-Clausal ALC θ-Connection Calculus) Figure $7$ shows the rules of the formal non-clausal ALC θ-connection calculus. Rules are applied bottom-up. The words of the calculus are tuples $C, M, Path$, where $C$ is a clause, $M$ is a matrix corresponding to query $O \models \alpha$ and $Path$ is a set of literals. $C$ is called the subgoal clause. $C_1$, $C_2$ and $C_3$ are clauses. The index $\mu \in \mathbb{N}$ of a clause $C^\mu$ denotes that $C^\mu$ is the $\mu$-th copy of clause $C$, increased when Copy is applied for that clause (the variable $x$ in $C^\mu$ is denoted $x_{\mu}$). When Copy is used, it has to be followed by the application of Extension or Reduction, to avoid non-determinism in the rules application. The Blocking Condition is defined as follows: the new individual $x_{\mu}$ (if it is new, then $x_{\mu} \not\in N_O$, as in the condition) is only created if the set of concepts of the previously created individual $\tau(x_{\mu-1})$ is not a subset of the set of concepts of the penultimate copied individual, i.e., $\tau(x_{\mu-1}) \not\subseteq \tau(x_{\mu-2})$.

The calculus consists of six rules. The Axiom, Start, Reduction and Copy rules are the same as the ones from the ALC θ-Connection Calculus. The Extension rule was modified to contain a β-Clause and the Decomposition rule [9] splits subgoal clauses into their sub-clauses.

Lemma 1. (Matrix characterization). A matrix $M$ is valid if there exist an index $\mu$, a set of $\theta$-substitutions $\theta_i$ and a set of connections $S$, s.t. every path through $M^\mu$, the matrix with copied clauses, contains a $\theta$-complementary connection $L_1^\mu, L_2^\mu$ in $S$, i.e. a connection with $\theta(L_1) = \theta(L_2)$. The tuple $\langle \mu, \langle \theta_i \rangle, S \rangle$ is called a matrix proof.

Example 6. (Non-Clausal ALC θ-Connection Calculus). Figure $2$ shows the proof for the $F_1$ of Example $7$ using the matrix representation.

The proof starts (1) by choosing a clause from the consequent as the start clause, in this case, $\{\text{CatOwner}(a)\}$, and a literal of that clause is selected, $\text{CatOwner}(a)^0$. This literal is connected to
Converting $\mathcal{ALC}$ Connection Proofs into $\mathcal{ALC}$ Sequents

Axiom$(A)$ \hspace{1cm} \{ \}, M, Path

Start$(S)$ \hspace{1cm} \begin{align*} & C_1, M, \{ \} \quad \text{with } C_1 \in \alpha \\ & e, M, e \end{align*}

Reduction$(R)$ \hspace{1cm} \begin{align*} & C, M, Path \cup \{ L_2 \} \\ & C \cup \{ L_1 \}, M, Path \cup \{ L_2 \} \text{ with } \theta(L_1) = \theta(T_2) \text{ and the Skolem condition holds} \\ & C_2 \text{ is an extension clause of } M \text{ wrt. } Path \cup \{ L_1 \}, L_2 \in C_2, \theta(L_1) = \theta(T_2) \text{ and the Skolem condition holds} \\ & C_3, M, Path \cup \{ L_1 \} \quad C, M, Path \text{ with } C_3 := \beta-\text{clause}_{L_1}(C_2), \end{align*}

Extension$(E)$ \hspace{1cm} \begin{align*} & C_3, M, Path \cup \{ L_1 \} \\ & C \cup \{ L_1 \}, M, Path \text{ with } \theta(L_1) = \theta(T_2) \text{ and the Skolem condition holds} \\ & C_2 \text{ is a copy of } C_1, L_2 \in C_2, \theta(L_1) = \theta(T_2) \text{ and the blocking condition holds} \end{align*}

Decomposition$(D)$ \hspace{1cm} \begin{align*} & C \cup C_1, M, Path \text{ with } C_1 \in M_1 \\ & C \cup \{ M_1 \}, M, Path \text{ with } \theta(L_1) = \theta(T_2) \text{ and the Skolem condition holds} \end{align*}

Copy$(C)$ \hspace{1cm} \begin{align*} & C \cup \{ L_1 \}, M \cup \{ C_2^w \}, Path \\ & C \cup \{ L_1 \}, M, Path \text{ with } C_2^w \text{ is a copy of } C_1, L_2 \in C_2^w, \theta(L_1) = \theta(T_2) \text{ and the blocking condition holds} \end{align*}

Figure 1: Non-clausal $\mathcal{ALC}$ $\theta$-Connection Calculus.

$CatOwner^1$ by an extension step and instance $a$ is the $\theta$-substitution of $CatOwner^1$ and $CatOwner(a)^0$. This connection is still not enough to prove all the paths starting from $CatOwner(a)^0$; the paths that start in it and pass through the literals from the other connected clause, namely, $Cat^0$ and $hasPet^0$, are still to be verified. Indeed, each connection creates two sets of literals to be checked, the remaining literals from each of the clauses involved in the connection. In the new extension step (2), the connection $\{Cat^0, Cat^1\}$ is established on the variable (or fictitious individual) $x$, as it is not necessary yet to commit the substitution with an already existing individual. There is still remaining literals to be verified, the ones resulting from the clause to which $Cat^0$ belongs. Next (3), the $hasPet^0$ predicate is connected, and the $\theta$-substitution generates the pair $(y,x)$ (not shown in figure), for the connection. $OldLady^0$ is connected to $OldLady(a)^1$ (4), and then (5), when the connection $\{hasPet^0, hasPet^1\}$ is settled (using a reduction step, as there was already a connection with the same literal in the path), $y$ was $\theta$-substituted by $y$ (i.e., $\theta(y) = a$), thus forming the pair $(a,x)$. This $\theta$-substitution over $y$ is then propagated through the path. Since every path through $M_1$ contains a $\theta$-complementary connection, $F_1$ is valid. However, the readability of the proof is largely lost by the transformations applied on the formulas to be proven, making it difficult to translate the steps into natural language.

Figure 2: The $\mathcal{ALC}$ non-clausal matrix proof of the $F_1$ using the graphical matrix representation.

Next, we present the Sequent Calculus to which $\mathcal{ALC}$ non-clausal proofs will be translated.
4 An ALC Sequent Calculus

According to [4], sequent calculi axiomatizes the relation of logical consequence (entailment), and this has an obvious parallel with the relation of subsumption, which is a keystone for DL representation and calculi. Bearing this in mind, Borgida et al proposed a sequent calculus for subsumption in-ferences in ALC as an extension of the standard sequent calculus, in which there are no rules of implication, as they are indeed subsumption rules, so implication is replaced by $\vdash$ without loss of meaning. In their calculus, terms are not moved from one side to the other of the turnstile during the proof, thus preserving the structure of the original subsumption, and in the case of multiple subsumptions, parentheses help in identifying the main subsumptions. Because of that, additional rules were created in which the negation is inserted in front of each construct, thus eliminating negation rules (l $\neg$, r $\neg$), what requires changing sequent antecedents to successors and vice versa. The calculus is divided in three parts: the first two describe sets of rules, while the last describes a set of axioms (see Figure 3, where $a$ and $b$ are arbitrary formulas and $X$ and $Y$ are arbitrary sequences of formulae).

- **Rules for propositional formulae:** rules $\sqcap$ and $\sqcup$ are duplicated by adding the negation rules for these connectives ($\neg\sqcap$, $\neg\sqcup$), while the proper negation rules ($\neg$) were modified to include the double negation rule ($\neg\neg$);

- **Rules for quantified formulae:** in [4], modal formulae are used (r $\square$, l $\Diamond$) and their negated rules (l $\neg\square$, r $\neg\Diamond$). Here, we replace these rules by their equivalents (r $\forall$, l $\exists$) and (l $\neg\forall$, r $\neg\exists$). The $\exists$-rules are the dual $\forall$-rules. A condition is explicitly considered for the application of these rules: the rule applies only if all homologous universal and existential formulae (e.g. $\forall h.C$ and $\exists h.C$ are homologous, $\forall h.C$ and $\exists f.C$ not) are joined together on the left and right sides of the sequent in the precondition. The rule is then applied only once;

- **Termination axioms:** unlike the standard sequent calculus, there are six termination axioms; all of them can be reduced to $X, a \vdash a, Y$ by applying the rules. The application of the $\neg$-rules forces formulae from the antecedent to the successor or vice versa, to be transformed until it gets to $X, a \vdash a, Y$, a procedure that is avoided in this calculus. Therefore, the additional termination axioms are necessary to ensure that formulae are never shifted from one side of the sequent to the other.

Although not stated explicitly, the calculus contains a cut rule, and the cut elimination theorem is valid in this case; it is stated below.

**Theorem 1. Cut Elimination Theorem [6].** Let $S$ be a set of sequents (axioms) and $s$ an individual sequent. $S \vdash SC s$, if and only if, there is a proof in SC of $s$ whose leaves are either logical or sequent axioms obtained by the substitution of $S$-belonging sequents, where the cut rule, $\Gamma, \Sigma \vdash \Delta, \Pi \frac{\Delta, \Sigma \vdash \Delta, \Pi}{\Gamma, \Sigma, \Delta \vdash \Pi}$, is only applied with a premise being an axiom.

**Example 7. (Sequent Proof for ALC Subsumption).** Figure 4 shows Example 7’s proof using the sequent calculus for ALC. The cut rule is applied to the initial assumptions, according to theorem 1. This proof tree could be described by the following text in natural language: (1) If individuals who own at least one cat as a pet are owners of cats; and if the old ladies are, individuals who have at least one animal as a pet and all individuals who have only cat as pet. So this implies that old ladies own cats. (2) So, the old ladies are all people who have at least one cat as a pet. And all individuals who own at least one cat as a pet, own cats. (3) In addition to old ladies are all individuals who have at least one animal as a pet and all individuals who have only cat as pets; all individuals who have at least one animal as a pet and all individuals who have only cat as pets, are all individuals who have at least one cat as a pet. (4) Thus, an animal or a cat implies in a cat.
Rules for propositional formulae

\[
\frac{X, a, b \vdash Y}{X, a \cup b \vdash Y} \quad (\text{l}\mathfrak{r}) \quad \frac{X \vdash a, Y \quad X \vdash b, Y}{X \vdash a \cup b, Y} \quad (\text{r}\mathfrak{r})
\]

\[
\frac{X, \neg a \vdash Y \quad X, \neg b \vdash Y}{X, \neg(a \cup b) \vdash Y} \quad (\text{l}\neg\neg) \quad \frac{X \vdash \neg a, b, Y}{X \vdash \neg(a \cup b), Y} \quad (\text{r}\neg\neg)
\]

\[
\frac{X, a \vdash Y \quad X, b \vdash Y}{X, a \cup b \vdash Y} \quad (\text{l}\mathfrak{i}) \quad \frac{X \vdash a, b, Y}{X \vdash a \cup b, Y} \quad (\text{r}\mathfrak{i})
\]

\[
\frac{X, \neg a, \neg b \vdash Y}{X, \neg(a \cup b) \vdash Y} \quad (\text{l}\neg\neg) \quad \frac{X \vdash \neg a, b, Y}{X \vdash \neg(a \cup b), Y} \quad (\text{r}\neg\neg)
\]

\[
\frac{X, a \vdash Y}{X, \neg \neg a \vdash Y} \quad (\text{l}\neg\neg) \quad \frac{X \vdash a, Y}{X \vdash \neg \neg a, Y} \quad (\text{r}\neg\neg)
\]

Rules for quantified formulae

\[
\frac{X' \vdash b, Y'}{X \vdash \forall r, b, Y} \quad (\text{r}\forall) \quad \frac{X, b \vdash Y'}{X, \exists r, b \vdash Y} \quad (\exists)
\]

\[
\frac{X', \neg b \vdash Y'}{X, \neg \forall r, b \vdash Y} \quad (\text{l}\forall) \quad \frac{X \vdash \neg b, Y'}{X \vdash \neg \exists r, b, Y} \quad (\text{r}\exists)
\]

where \(X' = \{a \mid \forall r, a \in X\} \cup \{\neg a \mid \exists r, a \in X\}\), and

\(Y' = \{a \mid \exists r, a \in Y\} \cup \{\neg a \mid \forall r, a \in Y\}\)

Termination axioms

\[
\frac{X, a \vdash a, Y}{=} \quad \frac{X, \neg a \vdash \neg a, Y}{=} \quad \frac{X, a \vdash \neg a, Y}{(\mathfrak{t}\forall)} \quad \frac{X \vdash a, \neg a, Y}{(\mathfrak{t}\mathfrak{r})}
\]

\[
\frac{X, \bot \vdash Y}{(\mathfrak{l}\bot)} \quad \frac{X \vdash \top, Y}{(\mathfrak{l}\mathfrak{t})}
\]

Cut rule

\[
\frac{\Gamma \vdash \Delta, A \quad A, \Sigma \vdash \Pi}{\Gamma, \Sigma \vdash \Delta, \Pi}
\]

Figure 3: The Sequent Calculus for \(\mathcal{ALC}\) Subsumption [4].

\[
\begin{align*}
\text{TRUE} & \quad \vdash A, C \cup C \quad \text{[3]} \\
\text{OL} + \exists h, A \cap \forall h, C & \quad \vdash \exists h, A, \forall h, C \quad \text{[1\text{\parallel}]} \\
\text{OL} + \exists h, C & \quad \vdash \exists h, A, \forall h, C + \exists h, C \quad \text{cut} \\
\text{OL} + \exists h, C + \exists h, C \quad \vdash \exists h, C + CO \quad \text{cut} \\
\text{OL} + \exists h, A \cap \forall h, C + \exists h, C \quad \vdash \exists h, C + CO \quad \text{cut} \\
\text{((C \cap CO) \cup (OL + \exists h, A \cap \forall h, C)) + (OL + CO)} \quad \vdash \exists h, C + CO \quad \text{[1\text{\parallel}]} 
\end{align*}
\]

Figure 4: \(\mathcal{ALC}\) sequent proof for \(F_1\). The names of the clauses and the roles are abbreviated.

5 Conversion Method

The process consists of two steps: building a formula tree and then converting this formula tree into sequents, given an \(\mathcal{ALC}\) query and its matrix non-clausal connection proof. They are explained below.

5.1 Building the Formula Tree

Definition 14. (Formula Tree, Position, Label, Polarity, Type). A formula tree is a syntactic representation of a formula \(F\) as a tree, where each node can have up to two child nodes. Each node has:

- **Position**: an index that identifies each element (predicate or connective) in the formula. Its represented as \(a_0, a_1, a_2, \ldots\);
- **Label**: either a connective (\(\land, \lor, \neg, \exists, =\)), quantifier or predicate, if it is an atomic
(sub-)formula. Nodes whose label is a predicate are leaves of the tree (figure 5b), while other nodes are internal (figure 5a): Polarity: can be 0 or 1. It is determined by the label and the parent node polarity. The root node of the tree has polarity 0; Type: the type of a node is a Greek letter: \( \alpha, \beta, \alpha', \beta', \gamma \) and \( \delta \). It is determined by its label and its polarity. Leaf nodes have no type. The polarity and type of a node are defined in table 2. For example, in the first line of this table, \((A \sqcap B)^1\) means that the node labelled \( \sqcap \) and polarity 1 has type \( \alpha \) and its successor nodes have polarity 1.

![Figure 5: Node Representation.](image)

<table>
<thead>
<tr>
<th>Type</th>
<th>Label (position/instances)</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>A(^1) B(^1)</td>
<td>( \alpha )</td>
</tr>
<tr>
<td>( \beta )</td>
<td>A(^0) B(^0)</td>
<td>( \beta )</td>
</tr>
<tr>
<td>( \delta )</td>
<td>(( \forall r ) A(^r))</td>
<td>( \delta )</td>
</tr>
<tr>
<td>( \alpha' )</td>
<td>A(^1)</td>
<td>( \alpha' )</td>
</tr>
<tr>
<td>( \beta' )</td>
<td>A(^1) B(^1)</td>
<td>( \beta' )</td>
</tr>
<tr>
<td>( \gamma )</td>
<td>(( \forall r ) A(^r))</td>
<td>( \gamma )</td>
</tr>
</tbody>
</table>

Nodes of type \( \alpha \) and \( \alpha' \) correspond to sequent rules that do not cause proof branching. Nodes of type \( \gamma \) and \( \delta \) correspond to quantifier rules. Rules associated to type \( \delta \) have the eigenvariable condition in the sequent calculi (where the term \( t \), the eigenvariable in the inference, appears in the main formula of inference and in no other formula in the sequent. In the case of the \( \exists \) rule for the existential quantifier and \( \forall \) rule for the universal quantifier). Nodes of type \( \beta \) and \( \beta' \) (i.e., \( \sqcap^0, \sqcup^1 \), and \( \sqsubseteq^1 \)) are particularly important, since their respective rules in sequents (described in table 2) split proof branching into two independent sub-proofs. Nodes have their types indexed in the formula tree to facilitate their identification, for example \( \beta_1, \beta_2, \beta'_1, \beta'_2 \). Each branch whose root is of type \( \beta \) or \( \beta' \) is marked with a letter (a,b,c,....).

Leaf nodes with instances are children of nodes type \( \alpha, \alpha' \) or \( \beta \). Leaf nodes without instances have labels attached to their closest predecessor nodes’ position, according to the following criteria: (1) if the leaf node label represents a concept, it has an unique position associated to its label; (2) if the leaf node label represents a role, it has two positions associated to its label in the form \((a_1, a_2)\), where \( a_2 \) is the of the nearest predecessor node’s position; (3) only type \( \gamma, \delta \) and \( \beta' \) node positions are associated to the labels. This helps to check for complementarity in a connection between two nodes.

The tree construction is guided by the identification of the (sub-)formulae’s main constructor (connective or quantifier), which will be a label in the tree node. This node has at most two branches that binds them to their child nodes, i.e., new (sub-)formulas. The node type and its childrens polarities are assigned according to table 2. If children nodes are not atomic (sub-)formulae, the process repeats itself by identifying these (sub-)formulae’s main constructor and then generating other nodes in the tree, until it reaches the leaves.
The proof matrix elements must correspond to the leaf nodes in the formula tree, indicated by the position of the corresponding predicate, as explained in section 5.2 step 2.

**Example 8. (Building the Formula Tree Process).** Figure 6 shows the first step in the tree construction for \( F_1 \) from Example 7 ((\( \exists h.C \subseteq CO \) \( \cap \) (\( OL \subseteq \exists h.A \cap \forall h.C \))) \( \models \) (\( OL(a) \subseteq CO(a) \))). Its main constructor is \( \models \), the root node label, which, by definition has polarity 0; its position is \( a_0 \). According to table 2, its type is \( \alpha' \); its children nodes, on the right and left, have polarities 0 and 1, respectively, and both are sub-formulas of \( \models \) in \( F_1 \). This process continues until it reaches the leaf nodes, as shown in figure 7.

\[
(\exists h.C \subseteq CO) \cap (OL \subseteq \exists h.A \cap \forall h.C))^{1}\]

\[
\models^{0} a_0
\]

\[
(OL(a) \subseteq CO(a))^{0}
\]

**Figure 6: Step 01 Process of building the formula tree for \( F_1 \).**

**Figure 7: Formula Tree for \( F_1 \) with labels, polarities and types.**

For a given formula \( A, A', B, B', \Gamma \) and \( \Delta \) are used to denote the sets of node positions of type \( \alpha, \alpha', \beta, \beta', \gamma \), and \( \delta \), respectively.

**Definition 15. (Substitution of positions \( \sigma_\delta \), ordering relation \( \sqsubseteq_\delta \)).** It replaces positions of type \( \gamma \) for positions of type \( \delta \). A position substitution \( \sigma_\delta \) is a mapping of the set \( \Gamma \) of type \( \gamma \) node positions to the set \( \Delta \) of type \( \delta \) node positions. The \( \sigma_\delta \) substitution induces a partial ordering relation \( \sqsubseteq_\delta \) in \( \Delta \times \Gamma \) as follows: let \( u \in \Gamma \) and \( v \in \Delta \); if \( \sigma_\delta(u) = p \) then \( v \sqsubseteq_\delta u \) for all \( v \in \Delta \) occurring in position \( p \).

Since the sequent rules \( r\forall \) and \( \exists \) and their homologues \( l\forall \) and \( r\exists \) are restricted to the eigenvariable condition, the relation \( v \sqsubseteq_\delta u \) expresses that the node labelled by \( v \) must be reduced before reducing the one labelled by \( u \).

**Example 9. (Substitution of positions \( \sigma_\delta \), ordering relation \( \sqsubseteq_\delta \)).** Consider the formula tree in figure 7. Let \( u \) be the node labelled by \( \forall^1 \), with position \( a_{13} \) and type \( \gamma \), and let \( v \) be the node labelled by \( \exists^1 \), with position \( a_{10} \) and type \( \delta \). To replace the position of a Type \( \gamma \) node by the position of a type \( \delta \) node, It is necessary to reduce the type \( \delta \) node first, then the node with the position \( a_{10} \) must be reduced before the node with the position \( a_{13} \). Thus, for this example, the ordering relation \( \sqsubseteq_\delta \) is given by \( \exists^1 a_{10} \sqsubseteq_\delta \forall^1 a_{13} \), and the substitution \( \sigma_\delta(\forall^1 a_{13}) = a_{10} \). With this, we have \( \sigma_\delta = \{a_{13}/a_{10}\} \).
Definition 16. (Substitution of positions $\sigma_{\beta'}$). It replaces positions of type $\beta'$, $\gamma$, $\delta$ for instances or positions of type $\beta'$. Positions of the nodes of type $\beta'$, $\gamma$ and $\delta$, as well as instances, appear in atomic formulas, so a substitution of positions $\sigma_{\beta'}$ is a mapping of the set $\beta'/\gamma/\delta$ to instances or positions of nodes of type $\beta'$. Let $u$ be a leaf node with the positions of nodes of type $\beta'/\gamma/\delta$ associated to its label and $v \in \beta'$; if $\sigma_{\beta'}(u) = p$, where $p \in \beta'$ or $p$ is an instance.

Reducing a node means applying the sequent rule that corresponds to that node over a given (sub-)formula. Leaf nodes are not reduced.

Example 10. (Substitution of positions $\sigma_{\beta'}$). Consider the formula tree in Figure 7. Let $u$ be the node labelled by $OL^0$, with position $a_8$ and position $a_7$ of type $\beta'$ associated to its label, and let $v$ be the node labelled by $OL^1$, with position $a_{17}$ and instance $a$. The substitution for this in leaf $u$ in this case is $\sigma_{\beta'}(OL(a_7)^0) = a$. Therefore, $\sigma_{\beta'} = \{a_7/a\}$.

Definition 17. (Substitution $\sigma_{\text{Final}}$). It is a combination of $\sigma_\delta$ and $\sigma_{\beta'}$. A $\sigma_{\text{Final}}$ substitution consists of a substitution $\sigma_\delta$ and a substitution $\sigma_{\beta'}$, where $\sigma_{\text{Final}} := \sigma_\delta \cup \sigma_{\beta'}$.

Example 11. (Substitution $\sigma_{\text{Final}}$). Considering the two previous examples, $\sigma_{\text{Final}} = \{a_{13}/a_{10}, a_7/a\}$.

Definition 18. (Connection, $\sigma_{\text{Final}}$-complementary connection). A connection is a pair of leaf nodes labelled with the same predicate symbol and the same position associated with the label or the same instance, but with different polarities. If they are identical under $\sigma_{\text{Final}}$, the connection is a $\sigma_{\text{Final}}$-complementary connection.

Example 12. (Connection, $\sigma_{\text{Final}}$-complementary connection). Let the formula tree in Figure 7 be. The leaf nodes $h(a_2,a_3)^0$ and $h(a_7,a_{10})^1$ with positions $a_4$ and $a_{11}$, respectively, form a connection that is complementary under $\sigma_{\text{Final}} = \{a_2/a_7,a_3/a_{10}\}$.

Definition 19. (Tree Ordering $\prec$). The tree ordering $\prec$ of an F formula is the partial ordering of the nodes positions in the tree formula. $\prec$ is defined as follows: (i) the root occupies the smallest position with respect to this ordering, (ii) $a_i \prec a_j$ if and only if the position $a_i$ is below $a_j$ in the formula tree.

Example 13. (Tree Ordering $\prec$). In the tree from Figure 7 there are examples of tree ordering: $a_7 \prec a_9 < a_{13} < a_{15}$ and $a_0 < a_1 < a_2 < a_3$.

Definition 20. (Reduction Order $\triangleleft$). The transitive closure of the union of $\subseteq_\delta$, $\subseteq_{\beta'}$ and $\prec$ is called reduction order $\triangleleft$, i.e., $\triangleleft := (\prec \cup \subseteq_\delta \cup \subseteq_{\beta'})^+$.

Nodes $v \triangleleft u$ means that the node $v$ must be reduced before the node labelled by $u$ in the sequent proof. $\triangleleft$ determines the nodes’ reduction order, and helps determine which sequent rules are to be used and in which order.

Example 14. (Reduction Order $\triangleleft$). In Figure 7 the nodes with positions $a_7$, $a_{10}$, $a_{16}$ and $a_{13}$, have the following reduction order $\triangleleft$: (i) $a_7 \triangleleft a_{10}$; (ii) $a_7 < a_{13}$; (iii) $a_{10} \subseteq_\delta a_{13}$. The orderings’ union and the tree ordering determine the reduction order for these nodes: $a_7 \prec a_{10} \triangleleft a_{13}$.

Definition 21. (Final Admissible Substitution). An $\sigma_{\text{Final}}$ Substitution is admissible if the reduction order $\triangleleft$ is not reflexive. In this case, it is possible to construct a sequent proof.

A correspondence between node label, polarity and type with the sequent rules presented in section 4 is established in Table 3. Such correspondence is useful for the sequent proof construction, where the polarity helps in the identification of the rule. Polarity 1 represents a rule on the left (left or l); polarity 0, on the right (right or r), for cases where there is already an associated rule. For instance, in Table 3’s first line, for node $\bigwedge^1$ the rule is $l\bigwedge$, while for node $\bigwedge^0$ it is $r\bigwedge$. For cases where internal nodes are preceded by a node labelled by a negation, correspondences are in Table 3’s last four columns.
Table 3: Correspondence between label, polarity and type of a node, preceded or not by a node labelled with negation, to ALC Sequent rules.

<table>
<thead>
<tr>
<th>Type $\alpha$</th>
<th>Rule</th>
<th>Not preceded</th>
<th>Type $\beta$</th>
<th>Rule</th>
<th>Preceded</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\top$</td>
<td>$\top$</td>
<td>$\top$</td>
<td>$\top$</td>
<td>$\top$</td>
<td>$\top$</td>
</tr>
<tr>
<td>$\bot$</td>
<td>$\bot$</td>
<td>$\bot$</td>
<td>$\bot$</td>
<td>$\bot$</td>
<td>$\bot$</td>
</tr>
<tr>
<td>$\neg$</td>
<td>$\neg$</td>
<td>$\neg$</td>
<td>$\neg$</td>
<td>$\neg$</td>
<td>$\neg$</td>
</tr>
<tr>
<td>Type $\alpha'$</td>
<td>Rule</td>
<td>Type $\beta'$</td>
<td>Rule</td>
<td>Type $\gamma$</td>
<td>Rule</td>
</tr>
<tr>
<td>$\equiv$</td>
<td>$\equiv$</td>
<td>$\equiv$</td>
<td>$\equiv$</td>
<td>$\equiv$</td>
<td>$\equiv$</td>
</tr>
<tr>
<td>$\not\equiv$</td>
<td>$\not\equiv$</td>
<td>$\not\equiv$</td>
<td>$\not\equiv$</td>
<td>$\not\equiv$</td>
<td>$\not\equiv$</td>
</tr>
</tbody>
</table>

5.2 Conversion to Sequents

Given an ALC query and its matricial non-clausal connection proof, the conversion procedure transforms this proof into an ALC sequent proof. This process performs four steps, which are described below:

- **Step 1- Formula tree construction:** A syntactic representation in tree form is constructed for the input formula, containing nodes, as described in [14]. The position of each predicate is input to step 2, and the tree to steps 3 and 4. **Example:** The conversion process begins with the $F_1$ formula tree construction, described in definition [14] which resulted in the formula tree represented in figure [7].

- **Step 2- Matrix elements’ positions assignment:** Since proof matrix elements correspond to predicates in the formula and also to leaf nodes in the formula tree, this step assigns to each matrix element the position of the corresponding predicate. Its input is the matrix non-clausal connection proof and the position of predicates. Its output is input to step 3. **Example:** Each element of the matrix is assigned with the position of the corresponding predicate in the formula, see matrix in [8].

- **Step 3- (partial) sequent proof structure Construction:** The matrix non-clausal connection proof with the positions of each element and the formula tree are inputs for this step. To each matrix connection, the formula tree is examined in search for the leaf nodes that correspond to the connection. The paths between the root node and these nodes in the tree are analyzed to determine the order of nodes to be worked on and thus build a structure of the (partial) proof in sequents. This
structure provides information about the reduction order $<$, which helps determine the rules to be applied, and on the existence of the proof branch, given by the identification of the nodes of type $\beta$ and $\beta'$. The (partial) sequent proof structure constructed will be the input for step 4. **Example:** The first connection links element $CO(a)^0$, from position $a_{18}$, to element $CO^1$, of position $a_6$, which are complementary under the substitution $\sigma_{\beta'} = \{a_2/a\}$, see table [4]. The path between these leaf nodes is $\{a_{18}, a_{16}, a_0, a_1, a_2, a_6\}$. Since there is no ordering relation $\preceq_{\sigma}$ between the nodes of that path and there are two tree orderings given by $a_0 < a_{16}$ and $a_0 < a_1 < a_2$. It is possible to start with any of these tree orderings. Choosing the first, we have the order of reduction at that moment equal to: $a_0 < a_{16} < a_1 < a_2$. Since the node with position $a_2$ is of type $\beta'$, the sequent is divided into two branches, called $c$ and $b$, as in the formula tree. Thus, this connection closes the branch $b$, branch where the node $CO^1$ is, and leads to the axiom $h^0, CO^0 \vdash CO^1$, because nodes of type $\beta'$ are associated with the cut rule (see table [3]). In the second connection, $C^0$, with position $a_5$ in branch $a$, is connected to $C^1$, with position $a_{15}$ in branch $d$, and the path between them is $\{a_5, a_3, a_2, a_1, a_7, a_9, a_{13}, a_{15}\}$. As the nodes with positions $a_1$ and $a_2$ have already been reduced, it is necessary to reduce the nodes with positions, $a_3$, $a_7$, $a_9$ and $a_{13}$, which have tree ordering $a_7 < a_9 < a_{13}$ and the relations $a_{10} \subseteq a_3$ and $a_{10} \subseteq a_{13}$. At the moment it is only possible to reduce the node with position $a_7$ and then the node with position $a_9$, that is, $a_7 < a_9$. Since the node with position $a_7$ is of type $\beta'$, its reduction divides branch 'a' into branches 'c' and 'd'. Then the node with position $a_9$, in branch 'd', is reduced. Since there are pendant nodes on this path, it is not yet possible to form an axiom and close the 'd' branch. The third connection is analyzed, where $h^1$, with position $a_{14}$, is connected to $h^1$, with position $a_{11}$, both in branch 'd'. The path between the nodes is $\{a_{14}, a_{13}, a_9, a_{10}, a_{11}\}$. Since $a_9$ has already been reduced, and there are the relations $a_{10} \subseteq a_{13}$ and $a_{10} \subseteq a_3$, the $a_{10}$ position node is reduced, and 'together' with it the nodes with position $a_{13}$ and $a_3$. The reduction of the $a_{10}$ position node makes the third and second connection complementary under the substitutions $\sigma_{\delta} = \{a_{13}/a_{10}, a_3/a_{10}\}$. With this the last two connections are reflected in the sequent proof leading to the closure of the 'd' branch. Notice that the second connection was only reached in the tree after the third connection, this leads to the axiom in the form $C^1 \vdash C^0$. The fourth connection connects $OL^0$, with position $a_8$ in branch 'c', to $OL^1$, with position $a_{17}$. The path between the nodes with these positions is $\{a_8, a_7, a_1, a_0, a_{16}, a_{17}\}$. As all nodes on this path have already been reduced, no reduction will be necessary in this step. Thus, 'c' branch is closed with an axiom in the form $OL^0 \vdash h^1A^1, h^0C^1$, due to the cut rule. This connection is complementary under $\sigma_{\beta'} = \{a_7/a\}$. On the fifth and last connection, which connects $h^0$ to $h^1$, there is no need of node reduction, since all nodes in the path were reduced. The connection is complementary under $\sigma_{\delta} = \{a_3/a_{10}\}$. Note that $a_2/a$ and $a_{17}/a$ were $\sigma_{\beta'}$ previous substitutions. All connections are complementary under a substitution $\sigma_{\text{Final}}$, all branches of the proof structure in sequent were closed, and the reduction order is not reflexive, as shown in figure [8] and in Table [4].

- **Step 4- Construction of the complete sequent proof:** Here, the process builds a complete sequent proof (output) from the (partial) sequent proof structure and the correspondence between nodes and sequent rules, described in [5]. The input is (partial) sequent proof structure, the formula tree and $\mathcal{ALC}$ sequent rules. **Example:** The structure obtained in step 3 is traversed. The proof begins with the reduction of $a_1$ position node, since the first two tree nodes do not have associated rule, because they are of type $\alpha'$. Rule $l\Pi$ is applied. Then, the $a_2$ position node, with type $\beta'$, is reduced by means of the cut rule on the query $\alpha$, that is, on $(OL \vdash CO)$. The proof is divided into branches 'a' and 'b'. The 'b' branch is closed with the initial axiom $\exists h.C \vdash CO$, while branch 'a' is open, in which $OL \vdash \exists h.C$ must be proved. The next node is of position $a_7$, of type $\beta'$, and its reduction
Table 4: Relation between connections, substitutions and orderings

<table>
<thead>
<tr>
<th>N</th>
<th>Nodes</th>
<th>σ_δ</th>
<th>σ_β</th>
<th>⊏_δ</th>
<th>&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CO(a_2)_1 a_6, CO(a)_9 a_18</td>
<td>a_2/a</td>
<td></td>
<td>a_0 &lt; a_16 &lt; a_1 &lt; a_2</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>C(a_3)_0 a_5, C(a_13)_1 a_15</td>
<td>a_13/a_10, a_3/a_10</td>
<td>a_10 ⊏_δ a_3, a_10 ⊏_δ a_13</td>
<td>a_7 &lt; a_9</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>h(a_7, a_13)_1 a_14, h(a_7, a_10)_1 a_11</td>
<td>a_13/a_10</td>
<td></td>
<td>a_10</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>OL(a_7)_1 a_8, OL(a)_1 a_17</td>
<td>a_7/a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>h(a_2, a_3)_1 a_4, h(a_7, a_10)_1 a_11</td>
<td>a_3/a_10, a_2/a</td>
<td>a_0 &lt; a_16 &lt; a_1 &lt; a_2 &lt; a_7 &lt; a_9 &lt; a_10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| σ_Final = a_2/a, a_13/a_10, a_3/a_10, a_7/a | a_0 < a_16 < a_1 < a_2 < a_7 < a_9 < a_10 |

divides the branch 'a' into branches 'c' and 'd', by means of the application of a new cut rule on \( \text{OL} \vdash \exists h.C \). The 'c' branch is closed with the initial axiom \( \text{OL} \vdash \exists h.A \sqcap \forall h.C \), while the 'd' branch stays open. To close the 'd' branch, the \( a_9 \) position node is reduced with the \( \sqcap \forall \) rule, followed by the node with position \( a_{10} \), through rule \( \exists \). This ends the \( F_1 \) sequent proof, as shown in figure 9.

\[
\frac{\forall h.A \sqcap \exists h.C}{\exists h.A \forall h.C \vdash \exists h.C} \quad \text{cut}
\]

Figure 9: Complete proof in \( \mathcal{ALC} \) sequents for \( F_1 \).

6 Complexity

This section presents a very brief overview of the main algorithms for the conversion method with its complexities, according to the 4 steps seen in section 5.2. All the algorithms are demonstrated in [10]. Time complexities were analyzed according to the input size of each algorithm. For example, some algorithms receive an \( \mathcal{ALC} \) formula \( F \) as input, so the input size \( n \) represents the number of symbols of \( F \). Other algorithms accept an \( F \) proof matrix as input; in this case, the input size is the matrix number of symbols, including connections between literals. This input is represented by \( m \).

Figure 10 presents the main algorithms’ execution order. Lines with arrows indicate that the output of one algorithm is input to another. For example, the output from algorithm 02 (called \textit{convertsPostfix}) is conveyed as input for algorithm 03 (called \textit{buildTree} and 04 (called \textit{assignPosition}). The complexity of algorithm 05 (\textit{Search Connections}) is the highest among the algorithms: \( O(n^3) \), up to four iterations over structures based on the input size \( m \).

7 Conclusions

This work presents a method to convert Non-clausal \( \mathcal{ALC} \) connections proofs into more readable proofs. The approach consists in transforming these proofs into proofs in the \( \mathcal{ALC} \)-Sequent Calculus [4]. Hence,
this conversion assumes that the input formulae will always be in non-clausal form, i.e., without the need to transform these formulae into any normal form. A tree representation of formulae is used as a guide in this conversion and a sequent proof is created while the connection proof is traversed. This conversion must contribute to describe how the reasoners based on the ALC Connection Method summon their inferences and may facilitate the creation of natural language explanations, given the ease of converting sequents to texts. The evaluation of the main algorithms’ computational complexities demonstrates its practical feasibility, since they display polynomial complexity. In this perspective, the scientific contributions of this work should characterize the importance of the logical proofs, clarify the reasoning process and increase inferences’ readability, thus providing better user interaction with connection reasoners.

References


