Propositional logics complexity and the sub-formula property

Edward Hermann Haeusler
PUC-Rio
Rio de Janeiro, Brasil
hermann@inf.puc-rio.br

In 1979 Richard Statman proved, using proof-theory, that the purely implicational fragment of Intuitionistic Logic ($\mathbf{M}_{\rightarrow}$) is PSPACE-complete. He showed a polynomially bounded translation from full Intuitionistic Propositional Logic into its implicational fragment. By the PSPACE-completeness of S4, proved by Ladner, and the Gödel translation from S4 into Intuitionistic Logic, the PSPACE-completeness of $\mathbf{M}_{\rightarrow}$ is drawn. The sub-formula principle for a deductive system for a logic $\mathcal{L}$ states that whenever $\{\gamma_1, \ldots, \gamma_k\} \vdash_{\mathcal{L}} \alpha$ there is a proof in which each formula occurrence is either a sub-formula of $\alpha$ or of some of $\gamma$. In this work we extend Statman’s result and show that any propositional (possibly modal) structural logic satisfying a particular formulation of the sub-formula principle is in PSPACE. If the logic includes the minimal purely implicational logic then it is PSPACE-complete. As a consequence, EXPTIME-complete propositional logics, such as PDL and the common-knowledge epistemic logic with at least 2 agents satisfy this particular sub-formula principle, if and only if, PSPACE=EXPTIME. We also show how our technique can be used to prove that any finitely many-valued logic has the set of its tautologies in PSPACE.

1 Introduction

In [17], R. Statman showed a polynomial-time reduction from Intuitionistic Propositional Logic into its implicational fragment. This reduction proves that Purely Implicational Minimal Logic is PSPACE-complete. The methods that Statman uses in [17] are based on proof-theory and Natural Deduction in Prawitz Style. The sub-formula principle for a Natural Deduction system $\mathcal{N}_{\mathcal{L}}$ for a logic $\mathcal{L}$ states that whenever $\alpha$ is provable from $\Gamma$, in $\mathcal{L}$, there is a derivation of $\alpha$ from a set of assumptions $\{\delta_1, \ldots, \delta_k\} \subseteq \Gamma$ built up only with sub-formulas of $\alpha$ and/or $\{\delta_1, \ldots, \delta_k\}$. In this article we show that the Validity problem for any propositional logic $\mathcal{L}$, with a Natural Deduction system that satisfies the sub-formula principle, is in PSPACE. Besides that, if $\mathcal{L}$ includes the usual $\rightarrow$-rules then PSPACE-complete.

In [13], we can found a general approach to Natural Deduction that allows the definition of general introduction and elimination rules in a way that, any intuitionistic logical constant (sometimes called operator or connective) that is expressible in terms of these general rules is also expressible by means of the intuitionistic logical constants ($\bot, \rightarrow, \land, \neg$ and $\lor$). Precisely, in [13] it is shown that any constant $c$ determined by a set of introduction and elimination rules, in Natural Deduction style, is such that, there is a formula $F(A_1, \ldots, A_k)$ built up with constants from $\{\bot, \rightarrow, \land, \neg, \lor\}$, and, $c(A_1, \ldots, A_k)$ is provable, iff, $F(A_1, \ldots, A_k)$ is provable in intuitionistic propositional logic. In order to prove this statement, namely, the functional completeness of intuitionistic logical constants, some proof-theoretical assumptions on the relationship between elimination and introduction rules are considered. Precisely, these assumptions have to do with the inversion principle that roughly states that an elimination rule for a logical constant $c$ has to be a function of respective introduction rules of this $c$. This means that the elimination rule is only determined by the introduction rules and this inversion principle.

1A good discussion on the form of the inversion principle can be found in [13]
Propositional logic's complexity

Although strongly based on [13], the results shown here, do not explicitly need the inversion principle. We do use a general form of introduction and elimination rules proposed in [13] and after extended by Roy Dickhoff and Nissim Francez to the term general-elimination harmony, such that, there is only one \( c \)-elimination rule associated to the many \( c \)-introduction rules. As we see in Section 2, this general-elimination harmony assumption is not needed, indeed. In [14] there is a very good discussion on the Higher-Level rules, proposed in [15] as an alternative and extension of Prawitz [13], and the general-elimination harmony as well. To sum it up, our technique relies in the sub-formula property only, as it is stated in the Definition 4 in the following section. We have to mention that [1] also provides an approach that does not require a full harmony between introduction and elimination rules, but [1] does not have the same goal as ours, in providing a computational complexity analysis based on proof-theoretical arguments as ours.

2 Translating a propositional logic \( \mathcal{L} \) into \( \mathcal{M} \rightarrow \)

Consider a propositional logic \( \mathcal{L} \) with Natural Deduction system having introduction and elimination rules according to the following general schema. A \( c \)-introduction rule, as shown in Figure 1, derives \( c(\beta_1, \ldots, \beta_n, \phi_1^1, \ldots, \phi_j^1, \ldots, \phi_i^n, \gamma_1, \ldots, \gamma_m) \), discharging occurrences of the bracket formulas \([\phi_1^1], \ldots, [\phi_j^1] \) in the respective derivation of \( \beta_i \). For reasons of readability, we sometimes write \( c(\beta, \phi^1, \gamma_k) \) as a shortened (abbreviation) of \( c(\beta_1, \ldots, \beta_n, \phi_1^1, \ldots, \phi_j^1, \ldots, \phi_i^n, \gamma_1, \ldots, \gamma_m) \). An elimination rule, corresponding to the same operator \( c \) is shown in Figure 2 below. It is important to say that the discharging discipline is liberal, that is, when applying a rule, it is possible to discharge multiple assumptions including vacuous discharges. In this way, this general schema cannot be used with sub-structural logics like Relevance and Linear Logics. It is also important to note that each \( \phi_j^i \) determine the type of the formula that can be discharged, not an instance of formula. For example, when representing the \( \rightarrow \)-introduction rule, there is only one type \( \phi_1^1 \) and \( c(\beta_1, \phi_1^1) \) is \( \phi_1^1 \rightarrow \beta_1 \). The \( \rightarrow \)-intro rule is shown in Figure 3 below as an instance of the general schema.

\[
\begin{array}{c|c|c}
[\phi_1^1], \ldots, [\phi_j^1] & [\phi_1^n], \ldots, [\phi_j^n] \\
\beta_1 & \ldots \\
\hline \\
\hline \\
\hline
\end{array}
\]

Figure 1: \( c \)-introduction rule schema

\[
\begin{array}{c|c|c|c|c|c|c}
[\beta_1] & [\beta_2] \\
\ldots & \ldots \\
\hline \\
\hline \\
\hline
\hline
\hline
\hline
\end{array}
\]

\[
\begin{array}{c|c|c}
\hline \\
\hline \\
\hline
\hline
\hline
\hline
\end{array}
\]

Figure 2: \( c \)-elimination rules schemata

\[\text{2 Consider usual } \lor \text{-intro rules. Both have } \beta_1 \lor \beta_2 \text{ as conclusion, one has } \beta_1 \text{ as premise and the other has } \beta_2 \text{ as the only premise. According to our schema, } \beta_2 \text{ plays the role of } \gamma_1 \text{ in the first rule, while } \beta_1 \text{ plays this role in the second } \lor \text{-intro rule.}\]
In elimination rules (see Figure 2), $\phi_j^i, i = 1, j_i$, are called simple-minor premise, for each $j$. The premises $\chi$ are called discharging minor premises (d-minor premise). The premise $c(\beta_i, \phi_j^i)$ is the major premise of the elimination rule. There may be more than one introduction rule, but at most one elimination rule. This elimination rule has one discharging minor premise for each introduction rule. As a matter of illustration we show below how the usual $\rightarrow$-elim rule is seen as instance of a general elimination schema. In order to see that, remember that the usual (minimal) implication has only one introduction rule that can be easily seen as an instance of an introduction schema. The elimination schema consider this instance of introduction schema, resulting in the shown rule. The usual form of the $\rightarrow$-elim is obtained by proving $B$ instead of discharging it. For a more detailed explanation on these general rules, we recommend [13].

Besides intuitionistic logic, some modal classical logics have Natural Deduction systems that conforms with our schema. For example, the Natural Deduction system presented in [12], page 75, for the logics S4 and S5 can be viewed as instances of our schemata, considering additional provisos for ensuring the soundness of the respective system. The rule for the $\Diamond$-introduction has no proviso and is as follows.

$$A \rightarrow B \quad A \quad C$$

where: (1) For S4, in the proof of the minor premise $B$, $B$ only depends on formulas $\Box \alpha$ or $\neg \Diamond \alpha$, but the formula occurrences $A$ discharged by the rule (2) For S5, in the proof of the minor premise $B$, $B$ only depends on formulas $\Box \alpha, \Diamond \alpha, \neg \Diamond \alpha, \neg \Box \alpha$, but the formula occurrences $A$ discharged by the rule.
In abstract proof-theory, we would require that the principle of inversion holds concerning the elimination and introduction rules. This usually is among the requirements to have a normalization theorem holding for a Natural Deduction system. However, for our analysis, nothing is required but the sub-formula principle, defined in Definition 4.

A logical constant \( c \), also called operator, can be a propositional connective or a modality. A Natural Deduction system for a logic \( \mathcal{L} \) is a set of introduction and/or elimination rules for the operators of \( \mathcal{L} \). Given a Natural Deduction system \( \mathcal{N}_\mathcal{L} \) for a logic \( \mathcal{L} \), the usual notions of derivation and proof of a formula from a set of assumptions are considered in this article. It is worth observing the difference between formula and formula occurrence.

**Definition 1 (Linked formulas and sequences)** Let \( \mathcal{N}_\mathcal{L} \) be a Natural Deduction for \( \mathcal{L} \). Let \( \Pi \) be a derivation of a formula \( \alpha \) from a set of assumptions \( \Gamma \) in \( \mathcal{N}_\mathcal{L} \). Let \( \beta_1 \) and \( \beta_2 \) be two formula occurrences in \( \Pi \). We say that \( \beta_1 \) is linked to \( \beta_2 \), if and only if, there is a sequence \( \delta_0, \ldots, \delta_k \) of formula occurrences in \( \Pi \), such that: (1) \( \delta_0 \) is \( \beta_1 \) and \( \delta_k \) is \( \beta_2 \), and, for each \( i = 0, \ldots, k - 1 \); (2) \( \delta_i \) is major premise of an application of a \( c \)-elim rule, in \( \Pi \), and \( \delta_{i+1} \) is a formula occurrence discharged by this application of elimination rule, or; (3) \( \delta_i \) is a discharging minor premise of an application of a \( c \)-elim rule, in \( \Pi \), and \( \delta_{i+1} \) is its conclusion, or; (4) \( \delta_i \) is a premise of an application of a \( c \)-intro rule, in \( \Pi \), and \( \delta_{i+1} \) is its conclusion.

We call \( \delta_0, \ldots, \delta_k \) a linking-sequence in \( \Pi \). Observing the definition above, we note that a formula occurrence that is either a simple-minor premise of an application of a \( c \)-elim rule in \( \Pi \) or the conclusion of \( \Pi \) cannot be linked to any other formula occurrence in \( \Pi \) any more. Thus, the longest possible linking-sequences in any derivation \( \Pi \) have \( \delta_k \) as a simple-minor premise or the conclusion of \( \Pi \). Our definition of linking-sequence is similar, and generalizes, the definition of branches as stated in [12].

**Definition 2** Let \( \Pi \) be a derivation in \( \mathcal{N}_\mathcal{L} \). Consider a linking-sequence \( \delta_0, \ldots, \delta_k \) in \( \Pi \). We say that \( \delta_0, \ldots, \delta_k \) is a sub-formula linking-sequence, if and only if, for every \( i = 0, \ldots, k \), either \( \delta_i \) is sub-formula of \( \delta_{i+1} \), or, \( \delta_{i+1} \) is sub-formula of \( \delta_i \).

**Definition 3 (Derivations satisfying the Sub-Formula Property)** Let \( \mathcal{L} \) be a logic and \( \mathcal{N}_\mathcal{L} \) a Natural Deduction system for \( \mathcal{L} \). Let \( \Pi \) be a derivation of \( \alpha \) from \( \{\delta_1, \ldots, \delta_k\} \) in \( \mathcal{N}_\mathcal{L} \). We say that \( \Pi \) satisfy the sub-formula principle, if and only if, every linking-sequence in \( \Pi \) is a sub-formula linking-sequence and every formula occurring in this linking-sequence is either a sub-formula of \( \alpha \) or sub-formula of some of the formulas in \( \Gamma \).

**Definition 4 (Systems satisfying the Sub-Formula Property)** We say that a system \( \mathcal{N}_\mathcal{L} \) satisfy the sub-formula property, if and only if, for every derivation \( \Pi \) of \( \alpha \) from \( \Gamma \) in it, there is a derivation \( \Pi' \) of \( \alpha \) from \( \Gamma' \subseteq \Gamma \) satisfying the sub-formula property.

When a logic \( \mathcal{L} \) has a system \( \mathcal{N}_\mathcal{L} \) satisfying the sub-formula property, we say that the logic \( \mathcal{L} \) itself satisfies the sub-formula property. Thus, it is important to emphasize that to satisfy the sub-formula property means that the logic has a Natural Deduction system according the general rules shown here. This also means that sub-structural logics do not satisfy our notion of sub-formula property.

**Proposition 1** Let \( \Pi \) be a derivation satisfying the sub-formula principle. If \( \Pi' \) is a sub-derivation of \( \Pi \) then \( \Pi' \) satisfies the sub-formula principle too.
Proof of proposition. We observe that a sub-sequence of a linking-sequence that is a sub-formula linking-sequence is a sub-formula linking-sequence too. Since every linking-sequence of Π is a sub-formula linking sequence, then every linking sequence of Π' is a linking sequence too.

Q.E.D.

In what follows we define a translation * from \( L \) into \( M_{→} \), such that, \( α \) is provable in \( L \), if and only if, \( α^* \) is provable in \( M_{→} \). In order to define *, we need auxiliary functions \( M \) and \( α' \), defined in the sequel. In the following definition, \( p_ω \) is a notation used to uniquely identify the symbol \( p \) indexed by the string (word) \( ω \). It this way we ensure that \( p \) is the unique propositional symbol that is named by means of the word \( ω \), such that, \( p_ω \) and \( p_ω' \) are equal, if and only if, \( ω_1 \) and \( ω_2 \) are the same word (string). We remember that we call operator any propositional connective or modality. We consider that the logic \( L \) has a finite set \( \{ c_1, \ldots, c_k \} \) of operators. For reasons of readability, we do not consider the minimal implication \( → \) in this set, even in the case that the logic has it as one of its propositional connectives. As we will see, the implication is the only logical constant (operator) that is not translated. So, in what follows we consider only logics including the purely minimal implicational logic.

The propositional logics considered in this article have in their Natural Deduction systems the usual \( → \)-intro and \( → \)-elim rules.

The following definitions provide us axioms schemata concerning each \( e \)-introduction and/or \( e \)-elimination rule.

**Definition 5 (I-axiom)** Consider an introduction rule \( r \) for an operator \( e \) as shown in Figure 7. The \( I \)-axiom concerning this rule schema \( r \) instantiated to formulas \( β_1 \) and \( φ_1^i, \ldots, φ_j^i \), denoted by \( I(r, β_1, φ_1^i, \ldots, φ_j^i) \), is the following implicational formula:

\[
(φ_1^i → (\ldots → (φ_j^i → β_1))) → \ldots (φ_n^i → (\ldots → (φ_j^i → β_n))) \rightarrow p_{e(β_1, φ_1^i, \ldots, φ_n^i, β_j^i, \ldots, β_j^i, β_n)}
\]

**Definition 6 (E-axiom)** Consider a formula \( Χ \) of a logic \( L \) and an elimination rule \( r \) for \( e \), as shown right in Figure 2, instantiated to \( φ_1^i, β_1 \) and \( Χ \). The \( E \)-axiom concerning this rule schema \( r \) and \( Χ \), denoted by \( E(r, Χ, β_1, φ_1^i) \), is the following implicational formula:

\[
\phi_1^i → (\ldots (φ_n^i → (β_1 → Χ) → \ldots (β_n → Χ)) → (p_{e(β_1, φ_1^i, Χ)}))
\]

For example, considering the usual \( \lor \)-introduction (\( \lor \)-intro1 and \( \lor \)-intro2) and \( \lor \)-elimination (\( \lor \)-elim) natural deduction schemata, we have that: (1) \( I(\lor \text{-intro1}, β_1, β_2) = β_1 \rightarrow p_{β_1 β_2} \); (2) \( I(\lor \text{-intro2}, β_1, β_2) = β_2 \rightarrow p_{β_1 β_2} \), and; (3) \( E(\lor \text{-elim}, β_1, β_2, Χ) = (β_1 → Χ) → ((β_2 → Χ) → (p_{β_1 β_2} → Χ)) \).

**Definition 7 (Atomizing Operators)** The mapping \( M \) from the language of \( L \) into the one of \( M_{→} \) is defined inductively, as follows: **Atoms** \( M(p) = p \), if \( p \) is a propositional letter; **Implication** \( M(α_1 → α_2) = M(α_1) → M(α_2) \); **Operators** \( M(c_m(β_1, φ_1^i, Χ)) = p_{c_m(β_1, φ_1^i, Χ)} \), if \( c_m \) is an operator of \( L \).

The second clause in the definition above is only used when the language of \( L \) includes the minimal implication \( → \). Otherwise it is not used and the translation is also well-defined. In the following we define an auxiliary function that for each formula \( α \) yields the set of implicational formulas that express the “deductive meaning”\(^3\) of the elimination and introduction rules for each operator in \( L \).

---

\(^3\)This informal expression is made precise in the statement of Theorem 1.
Definition 8 (Axiomatizing Operators) Given a formula $\alpha$ in $\mathcal{L}$, the mapping $A^\alpha$ from the language of $\mathcal{L}$ into (finite) sets of formulas in the language of $\mathcal{M}$ is defined inductively, as follows.

- **Atoms** $A^\alpha(p) = \emptyset$;
- **Implication** $A^\alpha(\alpha_1 \rightarrow \alpha_2) = A^\alpha(\alpha_1) \cup A^\alpha(\alpha_2)$;
- **Operators** $A^\alpha(c_m(\beta_i, \phi^i, \gamma_k)) =$

  $\{ \iota(r, M(\beta_i), M(\phi^i), M(\gamma_k)) / r \text{ is a } c_m\text{-intro rule} \} \cup \{ \varepsilon(r, M(\chi), M(\beta_i), M(\phi^i)) / r \text{ is a } c_m\text{-elim rule and } \chi \in \text{sub}(\alpha) \}$

Lemma 1 Let $\mathcal{L}$ be a logic having a Natural Deduction system satisfying sub-formula property. Let $\Pi$ be a proof of $\alpha$ from $\Gamma$ in $\mathcal{L}$. There is a derivation $\Pi'$ of $M(\alpha)$ from $\Gamma'$ where $\Gamma' \subseteq M(\Gamma) \cup \bigcup_{\gamma \in \Gamma} A^\alpha(\gamma)$ in $\mathcal{M}$. We use the notation $A^\alpha(\Gamma)$ to denote $\bigcup_{\gamma \in \Gamma} A^\alpha(\gamma)$ and “$\Gamma, \alpha$” to denote $\Gamma \cup \{ \alpha \}$.

Proof of Lemma. Since $\mathcal{L}$ satisfies the sub-formula principle, we can consider that $\Pi$ is a derivation satisfying the sub-formula principle. The proof proceeds by induction on the size of this derivation. The basis is the derivation of $\alpha$ from $\alpha$ itself, and hence, $\Pi'$ is $M(\alpha)$ only. The inductive step is according the last rule applied in the derivation satisfying the sub-formula principle. There are only two (general) cases:

- **Last rule is a $c_m$-intro rule.** Then $\Pi$ is as following:

  $\begin{array}{c}
  [\phi^1_i], \ldots, [\phi^j_i] \\
  \Pi^1_i \\
  \beta_i \\
  \hline
  \end{array}
  \begin{array}{c}
  [\phi^1_i], \ldots, [\phi^j_i] \\
  \Pi^1_i \\
  \beta_i \\
  \hline
  \end{array}
  \begin{array}{c}
  [\phi^1_i], \ldots, [\phi^j_i] \\
  \Pi^1_i \\
  \beta_i \\
  \hline
  \end{array}
  \begin{array}{c}
  [\phi^1_i], \ldots, [\phi^j_i] \\
  \Pi^1_i \\
  \beta_i \\
  \hline
  \end{array}
  \begin{array}{c}
  [\phi^1_i], \ldots, [\phi^j_i] \\
  \Pi^1_i \\
  \beta_i \\
  \hline
  \end{array}$

  Where the last rule is $r$, a $c$-introduction rule. Consider $t(r, \beta_i, \phi^i, \gamma_k)$ the implicational formula related to $c$. By Proposition $\Pi^i$ satisfies the sub-formula principle, $i = 1, n$. By inductive hypothesis, for each $i = 1, n$, there is a derivation $\Pi^i_\prime$ satisfying the statement of the lemma. Each $\Pi^i_\prime$ is of the following form:

  $\begin{array}{c}
  [\text{M}(\phi^1_i)], \ldots, [\text{M}(\phi^j_i)] \\
  \Pi^i_\prime \\
  \text{M}(\beta_i) \\
  \hline
  \end{array}$

  Thus, for each $i$ we can build, from $\Pi^i_\prime$, a derivation as the following, which we will denote by $\Sigma_i$. 

\[
\begin{align*}
[\mathcal{M}(\phi_1^i)], \ldots, [\mathcal{M}(\phi_{ji}^j)] & \\
\Pi_i' & \\
\rightarrow i & \mathcal{M}(\beta_i) \\
\mathcal{M}(\phi_i^j) & \rightarrow \mathcal{M}(\beta_i) \\
\mathcal{M}(\phi_{ji}^j) & \rightarrow (\mathcal{M}(\phi_i^j) \rightarrow \mathcal{M}(\beta_i)) \\
\vdots & \\
(\mathcal{M}(\phi_i^j) & \rightarrow (\ldots \rightarrow (\mathcal{M}(\phi_i^j) \rightarrow \mathcal{M}(\beta_i)))))
\end{align*}
\]

Finally, the derivation below is a derivation in \(\mathcal{M}_\rightarrow\) of the form stated by the lemma, that is, of \(\mathcal{M}(c(\beta_1, \ldots, \beta_n, \phi_1^1, \ldots, \phi_j^1, \ldots, \phi_n^1, \gamma_1, \ldots, \gamma_m)) = p(c(\beta, \phi, \gamma))\). For reasons of space we abbreviate \((\ldots \rightarrow (\mathcal{M}(\phi_i^j) \rightarrow \mathcal{M}(\beta_i)))\) by \(B_i\).

\[
\begin{align*}
[\mathcal{M}(\phi_1^1)], \ldots, [\mathcal{M}(\phi_{ji}^j)] & \\
[\mathcal{M}(\phi_1^n)], \ldots, [\mathcal{M}(\phi_{ji}^n)] & \\
\Sigma_n & \\
B_n & \rightarrow p(c(\beta, \phi, \gamma))
\end{align*}
\]

The derivation above shows the existence of a derivation of the translated conclusion from the translated premises.

• Last rule is a \(c_m\)-elim rule. Then \(\Pi^*\) is as following:

\[
\begin{array}{c}
\Pi^*_{pm} \\
\Pi^*_{i,j} \\
\Pi^*_1 \\
\Pi^*_n \\
c(\beta_i, \phi^j) \\
\phi_i^j \\
\chi \\
\chi
\end{array}
\]

By Proposition \(\Pi^*_{pm}, \Pi^*_{i,j}\) and \(\Pi^*_i, \Pi^*_j, i = 1, n, j = 1, k\) satisfies the sub-formula principle. By inductive hypothesis, for each \(i\) (and \(j\)), there are derivations \(\Pi_i', \Pi_j'\) satisfying the
Propositional logics complexity

We denote the derivation above by \( \Sigma \). Using \( \Pi' \), for each \( i \), we have the following derivation.

\[
\begin{align*}
\Pi' &
\begin{array}{c}
\Sigma
\end{array} \\
\begin{array}{c}
\vdots
\end{array} \\
\vdots \\
\begin{array}{c}
\vdots
\end{array} \\
\begin{array}{c}
\vdots
\end{array} \\
\vdots
\end{align*}
\]

Then the derivation below, built using the above derivations, shows how to conclude the translation of the conclusion of \( \Pi' \) from the translation of its premises and the implicational introduction and elimination schemata. \( \varepsilon(r, \mathcal{M}(\chi), \mathcal{M}(\bar{\beta}_i), \phi_{i,j}) \) is the \( \varepsilon \)-axiom that implements the implicational elimination on the implicational fragment of minimal logic. We remind the reader that \( \mathcal{M}(\varepsilon(\bar{\beta}_i, \phi^j)) = p_{c(\bar{\beta}_i, \phi^j)} \). We can check and find out that the derivation satisfies the lemma.

\[
\begin{align*}
\begin{array}{c}
\vdots
\end{array} \\
\begin{array}{c}
\vdots
\end{array} \\
\vdots \\
\begin{array}{c}
\vdots
\end{array} \\
\vdots
\end{align*}
\]
Q.E.D.

For logics that have rules with a proviso, the respective implicational axioms \( t \) and \( e \) must reflect the corresponding conditions. Whenever we have these axioms we have Lemma 1 holding for the respective logics. For example, this can be done with the Natural Deduction rules provided in [12] for S4 and S5 and already discussed before in this article.

**Proposition 2** Let \( \mathcal{L} \) be a propositional logic satisfying the sub-formula principle. Consider the following translation \( \star \) from formulas of \( \mathcal{L} \) into formulas of \( \mathcal{M} \rightarrow \) : Let \( \mathcal{A}(\alpha) = \{ \varphi_1, \ldots, \varphi_k \} \) \( \alpha^* \) is defined as \( \varphi_1 \rightarrow (\varphi_2 \rightarrow \ldots (\varphi_k \rightarrow \mathcal{M}(\alpha))) \). Thus, \( \vdash_{\mathcal{L}} \alpha \) if and only if \( \vdash_{\mathcal{M} \rightarrow} \alpha^* \).

**Proof of proposition.** The proposition follows immediately from Lemma 1.

Q.E.D.

**Proposition 3** If \( \alpha \) is any propositional formula, then the number of sub-formulas of \( \alpha \) is polynomially bounded on the length of \( \alpha \).

**Proof.** Each sub-formula of \( \alpha \) is determined by the main connective of it. In \( \alpha \) there is at most one connective or logical constant by symbol position in \( \alpha \). Thus, the number of sub-formulas of \( \alpha \) is bounded by the length of \( \alpha \).

We can see that the size of \( \alpha^* \) is \( O(m^3) \), if \( m \) is the size of \( \alpha \). The formula \( \alpha^* \) depends on three choices of sub-formulas of \( \alpha \). This polynomial bound on the size of \( \alpha^* \) entails the following main conclusions of our article.

**Theorem 1** If \( \mathcal{L} \) satisfies the sub-formula principle then the problem of knowing whether \( \alpha \), a formula of \( \mathcal{L} \), is provable or not, is in PSPACE. If \( \mathcal{L} \) includes \( \mathcal{M} \rightarrow \) then this problem, also known as Validity, is PSPACE-complete.

**Proof of theorem.** From the PSPACE-completeness of provability in \( \mathcal{M} \rightarrow \) and the polynomial reduction of \( \mathcal{L} \) to \( \mathcal{M} \rightarrow \), we have that Validity in \( \mathcal{L} \), namely, knowing whether \( \alpha \), a formula of \( \mathcal{L} \), is provable or not, is in PSPACE. If \( \mathcal{L} \) includes \( \mathcal{M} \rightarrow \), and since \( \mathcal{M} \rightarrow \) is PSPACE-complete, then Validity is PSPACE-complete.

Q.E.D.

**Corollary 1** If \( \mathcal{L} \) satisfies the sub-formula principle and includes \( \mathcal{M} \rightarrow \) then the problem of knowing whether \( \alpha \), a formula of \( \mathcal{L} \), is invalid or not, is PSPACE-complete.

**Proof of corollary.** Since PSPACE is a deterministic class, we have that CoPSPACE=PSPACE. The problem stated in the corollary is in CoPSPACE, by definition.

Q.E.D.

3 A brief discussion on extending the results to finitely many-valued propositional logics

An stronger version of the statement of Theorem 1 can be used to prove that some finitely many-valued logics have Validity (cf. Theorem 1) in PSPACE. In [5] it is provided a general schema to define normalizable Natural Deduction systems for some finitely many-valued logics. The kind of Natural Deduction rules used in [5] are similar, in fact almost the same, to those shown here in this article. However, when
formalizing logics with more than 2 truth-values, auxiliary formulas are considered in order to perform the binary print, explained in sequel, of a truth-value. We briefly discuss here the main idea on an extension of Theorem 1 and how it could be used to prove that Validity/Provability in a finitely many-valued logic is in PSPACE, by adjusting the sub-formula principle defined in Definition 3. More details on the Natural Deduction systems for finitely many-valued logics as used here can be found in [5].

Instead of defining a rule for each of the $n > 2$ truth-values, it is used a method to reduce many-valued semantics to bivalent one. We use the three-valued logic $L_3$, due to Łukasiewicz to illustrate our discussion. See [4] for the original article (1920) on $L_3$ and [9] for a very good and concise presentation of many-valued logics.

The truth-values in $L_3$ are $\{0, i, 1\}$, where 0 and $i$ are undesignated values and 1 is the designated value. With the sake of providing some context on the meaning of these three values, we can say that the truth-value $i$ means indetermination or in numerical terms that it corresponds to $1/2$. Thus, a formula having $i$ as value should not be taken as a true formula. The truth-values 0 and $i$ in this case are considered as undesignated. For some many-valued logics, a formula is valid, iff, it yields designated value for each possible truth-value assignment. This is just the case with $L_3$. $L_3$ has the truth-table shown in Table 4 for the implication $\rightarrow$ and in Table 5 for the negation $\neg$. We explain what is a bitprint of a truth-value in the sequel. Consider a function $t(P)$ that yields 1 if $P$ is designated and 0 if it is undesignated. Using the formula $\phi(P)$, defined as $\neg P \rightarrow P$, it is possible to identify $i$ with the pair $\langle 0, 1 \rangle$, 1 with the pair $\langle 1, 1 \rangle$ and 0 with $\langle 0, 0 \rangle$, that is, each truth-value $x$ is identified by the pair $\langle t(x), \phi(x) \rangle$. In [10], the sequence $\langle P, \phi(P) \rangle$ that denotes each one of the values in $\{0, i, 1\}$ is called the bitprint of the respective truth-value. It is interesting to note that the pair $\langle 1, 0 \rangle$ has no truth-value associated to it. Besides that, it is worth observing that for a many-valued logic with $k$ truth-values, we have to find $\max int(\log_2(k)) - 1$ separating formulas for performing the role of $\phi(P)$. Using bitprints, the truth-table of the $\supset$ is as shown in Table 6. In this way the line of the truth-table saying that $P \supset Q$ yields 1, when $P$ is $i$ and $Q$ is 1, is coded as: $P$ is 0, $\phi(P)$ is 1, $Q$ is 1 and $\phi(Q)$ is 1. This line, the fourth line of the truth-table, is related to the Natural Deduction rule shown in Figure 8. The rule in Figure 8, very similar to an elimination rule in Natural Deduction, corresponds to the third line in Table 6. Because some pairs $\langle t(P), \phi(P) \rangle$ do not correspond to any truth-value, there is need of a Natural Deduction rule to take care of this. The rule that is related to the pair $\langle 1, 0 \rangle$ in $L_3$ is shown in Figure 9. Besides that, rules as shown in Figure 10 are need in order to have a complete system, for, in the general case, formulas of the form $\phi(X)$, may have $X$ as composed formulas. Figure 10 shows the case when $X$ is $P \rightarrow Q$. This is the case related to the first line of the truth-table in Figure 6.

Consider a general propositional connective $c$, that forms formulas of the form $c(A_1, \ldots, A_n)$. The rules that discharge $c(A_1, \ldots, A_n)$ are considered $c$-introduction rules. The rules that have $c(A_1, \ldots, A_n)$ as premise (also called major premise) are the $c$-elimination rules. The same is said about $\phi(c(A_1, \ldots, A_n))$, where $\phi$ is a separating formula. In [5] it is shown how to eliminate sequences that starts with $c$-intro rules and ends with $c$-elimination rules. Derivations having this kind of sequences are said to be non-normal. Normal derivations do not have these sequences, and, do not have $c$-rules, like the one shown in Figure 9, proving premise of $c$-elimination rules either. It is the case that any non-normal derivation of a formula $\alpha$ from a set of formulas $\Gamma$ can be effectively transformed in a normal derivation of $\alpha$ from $\Gamma' \subseteq \Gamma$. This is the statement of the normalization theorem for finitely many-valued logics proved in [5].

A branch in a derivation in a Natural Deduction system for a finitely many-valued logic, as those discussed in this section, is any sequence of formulas starting in an undischarged hypothesis and ending in the minor premise of an elimination rule. This is essentially what is stated in Definition 1. One of the

---

4This depends whether we are in an epistemic position, the former, or an ontic position, the later case.
properties of normal derivations is that for any branch starting with a formula \( A \) and ending with \( B \) in this derivation, the formulas that occur in it are either sub-formulas of \( A \), or \( B \), or both. This is one of the main features of normal proofs, in usual systems as Classical, Intuitionistic and Minimal logic. As was already mentioned in the introduction, normalization implies in the sub-formula property in many logics. A quick inspection on the form of the general Natural Deduction introduction and elimination rules as defined in Section 2 shows that for these rules the premises are either sub-formula of the conclusion or vice-versa. The discharged formulas of both, introduction and elimination are also sub-formula either of the major premise or of the conclusion. In fact, the sub-formula principle holds because of this feature.
just discussed. However, the format of the Natural Deduction rules for the finitely many-valued logics does not show this perfect relationship. In their case, the use of the separating formulas $\phi(X)$ disturbs this relationship based only on sub-formula occurrence in rules. In order to apply our technique to finitely many-valued logic, we will consider the sub-formula relationship through application of any formula in the set of separating formulas $\{\phi_1, \ldots, \phi_n\}$, so that the relationship between conclusions, premises and discharged formulas, in this case of Natural Deduction rule, includes the separating formulas. For example, $P$, the discharged formula, and $Q$, a minor premise, are sub-formulas of the major premise $P \supset Q$, of the rule in Figure 8. However, $\phi(P)$, a discharged formula, and $\phi(Q)$, other minor premise, are not sub-formulas. They are, in fact, images of sub-formulas of the major premise. So, we can say that the formulas occurrences in a Natural Deduction rule, as in Figure 8 are $\phi$-related. This is defined below.

**Definition 9** Consider a Natural Deduction rule in a system for finitely many-valued logic using a set $\Phi = \{\phi_1, \ldots, \phi_n\}$ of separating formulas. The relationship between premises, conclusions and discharged formulas under sub-formula of $\phi_i$, $i = 1, n$, image of a sub-formula is denoted by $\Phi$-sub-formula relationship.

It is a trivial fact that if $\Phi \subseteq \Phi'$ and $\alpha$ is $\Phi$-sub-formula of $\beta$, or vice-versa, then they are $\Phi'$-sub-
formula too. From the form of a branch beginning with $\beta_1$ and ending in $\beta_k$ in a normal derivation we can see that any formula in this branch is either a sub-formula of $\beta_1$ or of $\beta_k$, and for any $i = 1, n - 1$, $\beta_i$ is $\Phi$-sub-formula of $\beta_{i+1}$, or vice-versa. Using this, we can use the same polynomial simulation used for proving Theorem 1 for any finitely many-valued logic in $M_n$. It is worth noting that the conclusion and hypothesis of any normal derivation impose that the set of formulas that occur in it are either sub-formulas of the conclusion or of the hypothesis. Does not matter the fact that some of them are $\Phi$-images of some other sub-formula.

As a matter of illustration of the discussion in the last paragraph, we show the implicational schemata for the elimination, introduction and $\phi$-rules, including the $\phi$-rules related to no bitprint representation (as the one shown in Figure 9). For example, in the case of $L_3$, the implicational schema for the rule shown in Figure 7 is below, where $C$ is a sub-formula of either the conclusion $\alpha$ of the derivation, or of some hypothesis of the derivation. There must be one such formula for each $C$, this already includes the images under the separating formula $\phi$. We remember that $p_{(P \supset Q)}$ is the new (fresh) propositional variable associated to $P \supset Q$. The same holds for the propositional variables $p_{\phi(P)}$ and $p_{\phi(Q)}$.

\[
(P \rightarrow C) \rightarrow ((P \rightarrow C) \rightarrow (p_{\phi(P)} \rightarrow (Q \rightarrow (p_{\phi(Q)} \rightarrow C))))
\]

The formula below is the implicational schema for the corresponding $\phi$-introduction rule in Figure 10.

\[
(P \rightarrow C) \rightarrow ((p_{\phi(P)} \rightarrow C) \rightarrow ((Q \rightarrow C) \rightarrow ((p_{\phi(Q)} \rightarrow C) \rightarrow ((p_{(P \supset Q)} \rightarrow C) \rightarrow C))))
\]

The implicational schema for the $\supset$-elimination rule (Figure 8 and the rule shown in Figure 10) are show in the sequel.

\[
(P \rightarrow C) \rightarrow ((p_{\phi(P)} \rightarrow C) \rightarrow (Q \rightarrow (p_{\phi(Q)} \rightarrow (p_{(P \supset Q)} \rightarrow C))))
\]

\[
P \rightarrow ((p_{\phi(P)} \rightarrow C) \rightarrow C)
\]

**Important Observation:** The reader must have noted that the forms of the implicational schemata are not totally in formal agreement with the $\iota$-axioms presented in the previous sections and used to prove that Validity of the respective logic is in PSPACE. This is only a formal disagreement, indeed. This can be seen, if we consider the usual conjunction $\land$. The $\land$ very well-known introduction rules is the left rule below, but it can be also defined as drawn by the right rule.

\[
\begin{array}{c|c|c}
A \land B \\
\hline
A & B & C \\
\hline
A \land B & A & B \\
\end{array}
\]

Finally, from the proof-theoretical results obtained in [5] we can state the following theorem. A detailed proof of this theorem and a more systematic treatment of the subject discussed in this section is omitted due to the lack of space.

**Theorem 2** Let $\mathcal{L}$ be a propositional $k$-valued logic with finitely-many connectives with semantics determined by functional truth-tables. The Validity decision problem for $\mathcal{L}$ is in PSPACE.
**Proof.** The fact that the connectives are defined by functional truth-tables, i.e. deterministic and defined for every \( k \)-value truth-tables, entails that the Natural Deduction system is sound and complete regarding the semantics provided by the truth-tables. Consider the set \( \{ \phi_1, \ldots, \phi_{log_2(k)} \} \) of separating formulas that provide bitprints for each of the \( k \) truth-values. There is at most one rule for each line in each truth-value. Since the number of lines in each truth-table is fixed (depends only on \( k \)), there is at most one implicational schema for each line. Let \( F(k) \) be this number. \( F(k) \) depends only on \( k \). In fact \( F(k) \) is induced by the arity of each connective of the language of \( \mathcal{L} \). Taking into account a formula \( \alpha \) in the language of \( \mathcal{L} \). Each implicational schema depends also on the number of sub-formulas of \( \alpha \), including those sub-formulas that are image under each \( \phi \). Let us say that \( n \) is the number of sub-formulas of the formula \( \alpha \), that is linear on the size of \( \alpha \). Thus, the total number of implicational schemata is at most \( n^{a \cdot \log_2(k)} \), where \( a = F(k) \). This upper-bound is polynomial in \( n \). Thus, the translation used in Theorem [1] can the used here to prove that Validity is in PSPACE. It is worth noting that \( F(k) \) depends only on the logic.

Q.E.D.

If the Validity of a logic \( \mathcal{L} \) is in PSPACE and it includes \( \mathbf{M} \rightarrow \) then its (Validity decision problem) is PSPACE-complete. A logic \( \mathcal{L} \) includes \( \mathbf{M} \rightarrow \), iff, the \( \rightarrow \)-intro and \( \rightarrow \)-elim are derived rules in \( \mathcal{L} \).

### 4 Conclusion

We have shown that structural propositional logics satisfying the sub-formula principle have their provability problem in PSPACE and that if they include \( \mathbf{M} \rightarrow \) then they are PSPACE-complete problems. We used proof-theory to show this. An important consequence of our result is to show that some logics hardly satisfy the sub-formula principle in the terms here presented. Any propositional logic that we believe that it is beyond PSPACE cannot satisfy the sub-formula principle.

An immediate application of Theorem [1] shows that hardly some well-known logics, as Propositional Dynamic Logic (PDL, see [6]) and the common-knowledge epistemic logic with at least 2 agents [7], have Natural Deduction systems satisfying the sub-formula principle. In [8], for example, it is shown a Natural Deduction system for PDL, with normalization theorem. However, one of its rules, the iteration rule, does not satisfy the sub-formula principle in the terms stated in this article, since it is an infinitary ND-system. Anyway it does not satisfy sub-formula relationship between premises and conclusion either. Any logic that is EXPTIME-complete, as the ones just discussed, cannot satisfy the sub-formula principle too.

Any PSPACE-complete problem is NP-Hard. Since Classical propositional Logic can be specified using the general elimination and introduction rules presented here, then it is PSPACE-HARD and so its also NP-Hard. We know, by the truth-table method that Classical propositional logic is in NP, so it is NP-complete.

We discussed the extension of this proof-theoretical way of providing upper-bounds for provability to finitely many-valued logics, concluding that their respective provability (validity) problems are in PSPACE. This is in our opinion the far we can go. The fact that the set of valid formulas of a logic is in PSPACE may be of no help in some particular analysis. There are many-valued logics, as for example Kleene’s logic, that have no tautologies and this is of course trivially in PSPACE. On the other hand, knowing that many-valued usual logics are in PSPACE seems to be an interesting fact in general. There are finitely many-valued logics that have non-trivial set of tautologies. As far as we know there should be no novelty in proving that Validity in finitely many-valued logics is in PSPACE. A polynomially space-bounded algorithm for testing Validity can be designed by observing that evaluating the truth-value of
a formula is polynomial on the size of the formula and that the evaluation of other rows can reuse the space. Here we presented an alternative proof of the fact that Validity is in PSPACE. The interesting part is to conclude that if the logic includes $M \rightarrow$, then its Validity is PSPACE-complete.

### 4.1 Further investigation

We have to consider the known PSPACE-complete propositional logics. The best well-known are the Modal Logics $K$, $S4$, and $KD$. We know that (unlabeled) Natural Deduction systems for these logics are problematic. They fail to satisfy normalization and have rules with complex provisos. A consequence of this is that the natural reductions expected to work according the inversion principle produce derivations out of the system. Thus, the application of our technique to these logics depends on a concrete Natural Deduction system with rules that are instances of the general rules we deal with. Anyway, we already know, by other methods that these logics are PSPACE-complete, so this application would not be worth of presenting. However, there are labeled Natural Deduction systems with normalization theorem and a related sub-formula lemma.

For logics that we do not know the complexity class of the provability problem, the extension of the technique presented in this article to the a kind of general labeled introduction and elimination system of rules would be worth of developing. Our general form for introduction and elimination rules does not consider the many approaches of labeled Natural Deduction. It is our intention to discuss how Theorem 1 can be obtained for the usual labeled systems used to specify modal logics, namely, [11], [16] and finally [2]. This would be quite useful in providing a proof-theoretical complexity analysis for the intuitionistic versions of Modal Logics.

Extending this technique to first-order logics would be interesting too. There are decidable fragments of pure predicate (without functions and the equality $\text{=}”) that are beyond PSPACE, and surely the statement of the sub-formula principle for these cases are worth of studying. Finally, investigations concerning the feasibility of the interpolant (see [3]) and the existence of ND provers with sub-formula property are worth of studying too.

*Acknowledgments:* We thank many colleagues for suggesting improvements in the work reported in this article. Particularly, Luiz Carlos Pereira, Lew Gordeev, Mario Benevides, Wagner Sanz and Peter Schroeder-Heister, Cecília Englander, Bruno Lopes and Jefferson de Barros Santos. We would like to mention Carlos Caleiro and João Marcos for pointing out that finitely many-valued logics could be a better example of positive application of our proof-theoretical technique than Modal Logics. We are very glad and thankful to have followed their suggestions.

### References


