A note on the size of prenex normal forms

Frederik Harwath[a]

[a] *Institut für Informatik, Goethe-Universität Frankfurt am Main, Postfach 11 19 32, 60054 Frankfurt am Main, Germany*

Abstract

The textbook method for converting a first-order logic formula to prenex normal form potentially leads to an exponential growth of the formula size, if the formula is allowed to use all of the classical logical connectives $\wedge, \vee, \to, \leftrightarrow, \neg$. This note presents a short proof which shows that the conversion is possible with polynomial growth of the formula size.

Keywords: algorithms, formal methods, logic in computer science, formula size and succinctness

1. Introduction

In algorithmic applications of logic, it is often helpful to assume that logical sentences have a special form, where all quantifiers occur in front of a formula which does not contain any further quantifiers. Formulae of this kind are called *prenex normal form* (pnf) formulae. The usual textbook proof (cf. e.g. [2],[3]) which shows that for each formula there is an equivalent pnf-formula leads to a simple algorithm. The formula constructed by this algorithm has the same size as the input formula if the input formula contains only the boolean connectives $\wedge, \vee, \neg$. However, classically, first-order sentences are often allowed to contain further connectives, in particular the implication $\to$ and the bi-implication $\leftrightarrow$. In this case, it is necessary to eliminate the further connectives using their definitions in terms of $\wedge, \vee, \neg$ before feeding the formula to the algorithm. Doing this naïvely for the bi-implication, i.e. $\varphi_1 \leftrightarrow \varphi_2$ is replaced by $(\varphi_1 \wedge \varphi_2) \vee (\neg\varphi_1 \wedge \neg\varphi_2)$, can lead to an exponential growth of the formula.

In the literature on computational logic, different reactions to this exponential growth can be encountered. Roughly speaking, it seems that either the exponential growth is simply accepted, or the admissible logical connectives are restricted, or all formulae are simply assumed to be in pnf. This situation seems rather unsatisfactory. Neither discussions with other researchers in computational logic nor a search of the literature answered the author's question whether the exponential growth is really necessary. It is the goal of this note to present a simple proof that it is indeed possible to efficiently convert formulae to pnf without restricting the admissible logical connectives. More precisely, we prove the following theorem.

Theorem 1. *Each first-order sentence φ over the base B_2 of all binary boolean connectives is equivalent to a sentence $\tilde{\varphi}$ over the base $\{\wedge, \vee, \neg\}$ of size*

$$\|\tilde{\varphi}\| \leq 4\|\varphi\|^{3.5}.$$

1.1. Notation

We assume only basic knowledge of first-order logic (see e.g. [2]). We continue with some general definitions which apply to both propositional formulae and to formulae of first-order logic. Let φ be such a logical formula. For a set B of boolean connectives, we say that φ is *over the base* B if all boolean connectives occurring in φ belong to B. The *syntax tree* $T(\varphi)$ of φ is defined as usual. We let $\ell(\varphi)$ denote the number of *leaves* of $T(\varphi)$. That is, if φ is a propositional formula, then $\ell(\varphi)$ counts the number of occurrences of variables in φ, and if φ is a first-order formula, then $\ell(\varphi)$ counts the number of occurrences of atomic subformulae. The *depth* of φ is the depth of $T(\varphi)$, i.e. the maximal number of edges on a directed path from the root to a leaf, and is denoted by $\mathrm{depth}(\varphi)$. The *quantifier-rank* of φ, written $\mathrm{qr}(\varphi)$, is the maximum number of quantifiers occurring on any directed path of $T(\varphi)$. We define the *size* $\|\varphi\|$ of φ as the number of nodes of $T(\varphi)$. Up to a constant factor (which, if φ is a first-order formula, depends on the signature of φ), $\|\varphi\|$ is the same as the length of φ as a word.

2. Reducing the base of first-order formulae

2.1. Quantifier-free formulae.

It is known that propositional logic formulae can be *balanced*, e.g. there is a constant k such that each propositional formula φ over the base $B = \{\wedge, \vee, \neg\}$ is equivalent to a formula $\tilde{\varphi}$ over the same base with $\mathrm{depth}(\tilde{\varphi}) \leq k \log(\ell(\varphi))$. (Here and throughout this note, log refers to the logarithm with base 2.) This result is usually attributed to Spira [5], but it has been proved independently several times; see [4] for an overview. The very same argument can be used to obtain a formula $\tilde{\varphi}$ over the base $\{\wedge, \vee, \neg\}$ if the original formula φ is over the base B_2 *of all binary boolean connectives.* In particular, $\|\tilde{\varphi}\|$ grows only polynomially, because $T(\varphi)$ is a binary tree and hence $\|\tilde{\varphi}\| \leq 2^{k \log(\ell(\varphi))+1} = 2\ell(\varphi)^k$.

Extending this result to first-order formulae would achieve our goal: we would first convert a first-order formula over the base B_2 to the base $\{\wedge, \vee, \neg\}$ and then we would use the standard algorithm to convert the resulting formula to pnf without further growth of the formula. Unfortunately, since the first-order quantifier-rank hierarchy is strict, it is impossible to reduce the depth of general first-order formulae in a similar way: consider the first-order sentence $\varphi_n := \exists x_1 \ldots \exists x_n \bigwedge_{1 \leq i < j \leq n} \neg(x_i = x_j)$ with $\ell(\varphi_n) = \frac{n(n-1)}{2}$ which states that there exist at least n distinct elements in a structure over empty the signature; it is well-known that each sentence $\tilde{\varphi}_n$ that is equivalent to φ_n must have quantifier-rank at least n and hence $\mathrm{depth}(\tilde{\varphi}_n) \geq n$. Nevertheless, we can use the result for propositional formulae to achieve our goal. First we note that the result about the balancing of propositional formulae translates to the following statement about first-order formulae.

Lemma 2. *Each quantifier-free first-order formula φ over the base B_2 is equivalent to a formula $\tilde{\varphi}$ over the base $\{\wedge, \vee, \neg\}$ with* $\mathrm{depth}(\tilde{\varphi}) \leq 2 \log_{\frac{3}{2}}(\ell(\varphi)) + 1$.

We could consider a quantifier-free first-order formula as a propositional logic formula and derive Lemma 2 from the balancing result for propositional formulae. To keep this note self-contained, we prefer to present its nice and short proof here. Our presentation borrows from [1]. Below, we write $\varphi \equiv \psi$ if φ and ψ are semantically equivalent.

Proof. We proceed by induction on $\ell := \ell(\varphi)$. If $\ell = 1$, let α be the sole atomic formula occurring in φ. By removing double negations, we see that φ is either equivalent to $\tilde{\varphi} := \alpha$ or to $\tilde{\varphi} := \neg\alpha$. In both cases, $\mathrm{depth}(\tilde{\varphi}) \leq 1$.

If $\ell \geq 2$, then φ contains a subformula ψ with $\lceil \frac{\ell}{3} \rceil \leq \ell(\psi) \leq \lfloor \frac{2\ell}{3} \rfloor$. To see why this is true, consider a subformula of ψ with $\ell(\psi) \geq \lceil \frac{\ell}{3} \rceil$ of minimal size. Towards a contradiction, suppose that $\ell(\psi) \geq \lfloor \frac{2\ell}{3} \rfloor + 1 \geq \lceil \frac{2\ell}{3} \rceil \geq 2$. Then ψ is not atomic and it has either one or two immediate subformulae, i.e. subformulae corresponding to children of the root of $T(\psi)$. For one such subformula ψ', we have $\ell(\psi') \geq \lceil \frac{\ell}{3} \rceil$ — a contradiction to the minimality of ψ.

Let φ_{true} and φ_{false} be the formulae obtained from φ by replacing an occurrence of the subformula ψ by atomic formulae true and false with the obvious meaning, respectively. Observe that

$$\varphi \;\equiv\; (\psi \wedge \varphi_{\text{true}}) \vee (\neg\psi \wedge \varphi_{\text{false}}).$$

We have removed at least $\lceil \frac{\ell}{3} \rceil \geq 1$ atoms, but introduced one new true- or false-atom. These new atoms can be eliminated, since e.g. $(\chi \wedge \text{true}) \equiv \chi$ and similar equivalences hold for all connectives. For the modified formulae, $\ell(\varphi_{\text{true}}), \ell(\varphi_{\text{false}}) \leq \ell - \lceil \frac{\ell}{3} \rceil \leq \lfloor \frac{2\ell}{3} \rfloor$. Now we apply the induction hypothesis to construct formulae $\tilde{\varphi}_{\text{true}}, \tilde{\varphi}_{\text{false}}$ and $\tilde{\psi}$, $\widetilde{(\neg\psi)}$. We let

$$\tilde{\varphi} \;:=\; (\tilde{\psi} \wedge \tilde{\varphi}_{\text{true}}) \vee (\widetilde{(\neg\psi)} \wedge \tilde{\varphi}_{\text{false}}).$$

Clearly, $\tilde{\varphi}$ is a formula over the base $\{\wedge, \vee, \neg\}$ and $\tilde{\varphi} \equiv \varphi$. Furthermore, we have $\mathrm{depth}(\tilde{\varphi}_{\text{true}}), \mathrm{depth}(\tilde{\varphi}_{\text{false}}), \mathrm{depth}(\tilde{\psi}), \mathrm{depth}(\widetilde{(\neg\psi)}) \leq 2\log_{\frac{3}{2}}(\lfloor \frac{2\ell}{3} \rfloor) + 1$. Hence,

$$\begin{aligned}
\mathrm{depth}(\tilde{\varphi}) \;&\leq\; \max\{\mathrm{depth}(\tilde{\varphi}_{\text{true}}), \mathrm{depth}(\tilde{\varphi}_{\text{false}}), \mathrm{depth}(\tilde{\psi}), \mathrm{depth}(\widetilde{(\neg\psi)})\} + 2 \\
&\leq 2\log_{\frac{3}{2}}(\tfrac{2\ell}{3}) + 3 \\
&= 2(\log_{\frac{3}{2}}(\ell) - 1) + 3 \;=\; 2\log_{\frac{3}{2}}(\ell) + 1.
\end{aligned}$$

□

2.2. Formulae with quantifiers

As observed above, we cannot hope for a similar depth reduction as in Lemma 2 for first-order formulae which contain quantifiers. To achieve our goal, we do not focus on the depth, but rather on the size of the formula; the depth of the formula in our construction below may even grow. Intuitively, the idea of the proof below is to perform the balancing "between the quantifiers".

Let $q(\varphi)$ denote the number of nodes of $T(\varphi)$ that are labelled by a quantifier (i.e. the number of occurrences of quantifiers in φ, not the quantifier-rank), and let $s(\varphi) := \ell(\varphi) + q(\varphi)$.

Theorem 3. *Each first-order formula φ over the base B_2 is equivalent to a formula $\tilde{\varphi}$ over the base $\{\wedge, \vee, \neg\}$ such that $\|\tilde{\varphi}\| \leq 4s(\varphi)^{3.5}$ and $\mathrm{depth}(\tilde{\varphi}) \leq 3.5(\mathrm{qr}(\varphi) + 1)(\log(\ell(\varphi)) + 1)$.*

Proof of Theorem 3. The proof is by induction on $q := q(\varphi)$. If $q = 0$, i.e. φ is quantifier-free, we obtain the formula $\tilde{\varphi}$ using Lemma 2. Note that $2\log_{\frac{3}{2}}(\ell) \leq 3.5\log(\ell)$. We have $\mathrm{depth}(\tilde{\varphi}) \leq 3.5\log(\ell(\varphi)) + 1$ and, since $T(\varphi)$ is a binary tree, $\|\tilde{\varphi}\| \leq 2^{3.5\log(\ell(\varphi))+2} \leq 4\ell(\varphi)^{3.5}$.

Now consider the case where $q \geq 1$. Let M denote the set of nodes v of $T(\varphi)$ such that v is labelled by a quantifier and such that no further quantifiers occur strictly above v. For the analysis of the formula size below, it will be important to distinguish between the nodes and the corresponding formulae. For each node $v \in M$, we let ψ_v denote the subformula of φ corresponding to the subtree of $T(\varphi)$ with root v. By definition, we have $\psi_v = Qx\ \psi'_v$ for some $Q \in \{\exists, \forall\}$. Let σ be the signature of φ and let $\sigma' := \sigma \cup \{R_{\psi_v} : v \in M\}$, where $R_{\psi_v} \notin \sigma$ is a new relation symbol whose arity equals the number of free variables of ψ_v. Let φ' be the quantifier-free formula over the signature σ' obtained from φ by replacing, for each $v \in M$ with $\psi_v = \psi_v(\bar{x})$, the subtree of $T(\varphi)$ with root v by a leaf labelled by the atom $R_{\psi_v}(\bar{x})$. Since, in φ', we have removed all atoms below each $v \in M$ and introduced one new atom for each $v \in M$, we have

$$\ell(\varphi) \;=\; \ell(\varphi') - |M| + \sum_{v\in M} \ell(\psi_v) \;=\; \ell(\varphi') - |M| + \sum_{v\in M} \ell(\psi'_v).$$

Furthermore, since $q(\varphi') = 0$ and, for each $v \in M$, $q(\psi'_v) = q(\psi_v) - 1$, we have

$$q(\varphi) \;=\; \sum_{v\in M} q(\psi_v) \;=\; \sum_{v\in M} (q(\psi'_v) + 1) \;=\; |M| + \sum_{v\in M} q(\psi'_v).$$

Hence,

$$s(\varphi) \;=\; \ell(\varphi) + q(\varphi) \;=\; \ell(\varphi') + \sum_{v\in M} s(\psi'_v).$$

By induction, since $q(\varphi') = 0 < q(\varphi)$, we obtain a formula $\tilde{\varphi}'$ over the base $\{\wedge, \vee, \neg\}$ and the signature σ' with $\|\tilde{\varphi}'\| \leq 4s(\varphi')^{3.5} = 4\ell(\varphi')^{3.5}$ and $\tilde{\varphi}' \equiv \varphi'$. Furthermore, for each $v \in M$, we have $q(\psi'_v) < q(\varphi)$, since $\psi_v = Qx\ \psi'_v$ for some $Q \in \{\exists, \forall\}$. Hence, we obtain a formula $\tilde{\psi}'_v$ over the base $\{\wedge, \vee, \neg\}$ of size $\|\tilde{\psi}'_v\| \leq 4s(\psi'_v)^{3.5}$ such that $\tilde{\psi}'_v \equiv \psi'_v$. We define

$$\hat{\psi}_v := Qx\ \tilde{\psi}'_v,$$

so that $\hat{\psi}_v \equiv \psi_v$. Let $\tilde{\varphi}$ be the formula over the signature σ obtained from $\tilde{\varphi}'$ by replacing each occurrence of an atom $R_{\psi_v}(\bar{x})$ with $v \in M$ by the formula $\hat{\psi}_v(\bar{x})$. To obtain $\tilde{\varphi}$, we have removed $|M|$ atomic formulae from φ. Furthermore, $\|\hat{\psi}_v\| = \|\tilde{\psi}'_v\| + 1$ for each $v \in M$. Hence, we obtain

$$\begin{aligned}
\|\tilde{\varphi}\| \;&\leq \|\tilde{\varphi}'\| - |M| + \sum_{v\in M} \|\hat{\psi}_v\| \\
&= \|\tilde{\varphi}'\| - |M| + \sum_{v\in M} (\|\tilde{\psi}'_v\| + 1) \\
&= \|\tilde{\varphi}'\| + \sum_{v\in M} \|\tilde{\psi}'_v\| \\
&\leq 4\big(\ell(\varphi')^{3.5} + \sum_{v\in M} s(\psi'_v)^{3.5}\big) \;\leq\; 4\big(\ell(\varphi') + \sum_{v\in M} s(\psi'_v)\big)^{3.5} \;=\; 4s(\varphi)^{3.5}.
\end{aligned}$$

Furthermore,

$$\begin{aligned}
\mathrm{depth}(\tilde{\varphi}) &\le \mathrm{depth}(\tilde{\varphi}') + \max_{v\in M} \mathrm{depth}(\hat{\psi}_v) - 1 \\
&= \mathrm{depth}(\tilde{\varphi}') + \max_{v\in M} \mathrm{depth}(\tilde{\psi}'_v) \\
&\le 3.5(\log(\ell(\varphi')) + 1) + \max_{v\in M} 3.5(\mathrm{qr}(\psi'_v) + 1)(\log(\ell(\psi'_v)) + 1) \\
&\le 3.5(\log(\ell(\varphi)) + 1) + 3.5\mathrm{qr}(\varphi)(\log(\ell(\varphi)) + 1) \\
&= 3.5(\mathrm{qr}(\varphi) + 1)(\log(\ell(\varphi)) + 1).
\end{aligned}$$

It remains to show that $\tilde{\varphi} \equiv \varphi$. For a σ-structure $\mathfrak{A}$, we let $\mathfrak{A}'$ be the σ'-expansion of $\mathfrak{A}$ with $R_{\psi_v}^{\mathfrak{A}'} := \psi_v(\mathfrak{A})$ for each $v \in M$. Here $R_{\psi_v}^{\mathfrak{A}'}$ denotes the interpretation of the relation symbol R_{ψ_v} in $\mathfrak{A}'$ and $\psi_v(\mathfrak{A})$ is the relation defined by ψ_v in $\mathfrak{A}$. This definition of $\mathfrak{A}'$ and the definition of φ' together with a straightforward structural induction show that $\mathfrak{A} \models \varphi$ iff $\mathfrak{A}' \models \varphi'$. Furthermore, $\tilde{\varphi}' \equiv \varphi'$, so that $\mathfrak{A} \models \varphi$ iff $\mathfrak{A}' \models \tilde{\varphi}'$. Considering the way we have replaced R_{ψ_v} by the formula $\tilde{\psi}_v$ to obtain $\tilde{\varphi}$ from $\tilde{\varphi}'$, we obtain that $\mathfrak{A}' \models \tilde{\varphi}'$ iff $\mathfrak{A} \models \tilde{\varphi}$, since for each $v \in M$, $R_{\psi_v}^{\mathfrak{A}'} = \psi_v(\mathfrak{A}) = \tilde{\psi}_v(\mathfrak{A}')$. Altogether, by transitivity, we have $\mathfrak{A} \models \varphi$ iff $\mathfrak{A} \models \tilde{\varphi}$. □

3. Conclusion

By the discussion above, it should be clear that our main result can be obtained as a consequence of Theorem 3.

Proof of Theorem 1. Apply Theorem 3 to convert φ to an equivalent formula $\tilde{\varphi}'$ over the base $\{\wedge, \vee, \neg\}$ of size $\|\tilde{\varphi}'\| \le 4s(\varphi)^{3.5} \le 4\|\varphi\|^{3.5}$. The usual algorithm (cf. e.g. [2], [3]) for converting a formula over the base $\{\wedge, \vee, \neg\}$ to pnf rewrites a formula according to equivalences between formulae of exactly the same size. Applying this algorithm to $\tilde{\varphi}'$ yields the desired formula $\tilde{\varphi}$. □

Note that Theorem 1 could be extended in several ways. The same kind of argument can be applied to bases containing boolean connectives of higher arity. The exponent in Theorem 1, which comes from the depth of the formula in Lemma 2, can be reduced; see [4] for an overview of related results, and in particular [1].

[1] Bonet, M., Buss, S. R., 1994. Size-depth tradeoffs for boolean formulae. Information Processing Letters 49 (3), 151 – 155.

[2] Ebbinghaus, H.-D., Flum, J., Thomas, W., 1994. Mathematical logic. Springer.

[3] Harrison, J., 2009. Handbook of Practical Logic and Automated Reasoning. Cambridge University Press.

[4] Jukna, S., 2012. Boolean Function Complexity - Advances and Frontiers. Vol. 27 of Algorithms and combinatorics. Springer.

[5] Spira, P., 1971. On time hardware complexity tradeoffs for boolean functions. In: Proceedings of the Fourth Hawaii International Symposium on System Sciences. pp. 525–527.