# Semantics of Attack-Defense Trees for Dynamic Countermeasures and a New Hierarchy of Star-free Languages 

Thomas Brihaye ${ }^{1[0000-0001-5763-3130]}$, Sophie Pinchinat ${ }^{2}$, and Alexandre Terefenko ${ }^{1,2}$<br>${ }^{1}$ University of Mons, Belgium<br>${ }^{2}$ University of Rennes, France


#### Abstract

We present a mathematical setting for attack-defense trees, a classic graphical model to specify attacks and countermeasures. We equip attack-defense trees with (trace) language semantics allowing to have an original dynamic interpretation of countermeasures. Interestingly, the expressiveness of attack-defense trees coincides with star-free languages, and the nested countermeasures impact the expressiveness of attack-defense trees. With an adequate notion of countermeasure-depth, we exhibit a strict hierarchy of the star-free languages that does not coincides with the classic one. Additionally, driven by the use of attackdefense trees in practice, we address the decision problems of trace membership and of non-emptiness, and study their computational complexities parameterized by the countermeasure-depth.


## 1 Introduction

Security is nowadays a subject of increasing attention as means to protect critical information resources from disclosure, theft or damage. The informal model of attack trees is due to Bruce Schneier ${ }^{3}$ to graphically represent and reason about possible threats one may use to attack a system. Attack trees have then been widespread in the industry and are advocated since the 2008 NATO report to govern the evaluation of the threat in risk analysis. The attack tree model has attracted the interest of the academic community in order to develop their mathematical theory together with formal methods (see the survey [24]).

Originally in [20], the model of attack tree aimed at describing how an attack goal refines into subgoals, by using two operators $O R$ and $A N D$ to coordinate those refinements. The subgoals are understood in a "static" manner in the sense that there is no notion of temporal precedence between them. Still, with this limited view, many analysis can be conducted (see for example [75]). Then, the academic community considered two extensions of attack trees. The first one, called attack-defense tree (adt, for short), is obtained by augmenting attack trees with nodes representing countermeasures [10]8]. The second one, initiated

[^0]by [16|6], concerns a "dynamic" view of attacks with the ability to specify that the subgoals must be achieved in a given order. This way to coordinate the subgoals is commonly specified by using operator $S A N D$ (for Sequential $A N D$ ). In [1], the authors proposed a path semantics for attack trees with respect to a given a transition system (a model of the real system). However, a unifying formal semantics amenable to the coexistence of both extensions of attack trees - namely with the defense and the dynamics - has not been investigated yet.

In this paper, we propose a formal language semantics of adts, in the spirit of the trace semantics by [2] (for defenseless attack trees), that allows countermeasure features via the new operator CO (for "countermeasure"). Interestingly, because in adts, countermeasures of countermeasures exist, we define the countermeasure-depth (maximum number of nested Co operators) and analyze its role in terms of expressiveness of the model.

First, we establish the Small Model Property for adts with countermeasuredepth bounded by one (Theorem 1), which ensure the existence of small traces in a non-empty semantics. This not so trivial result is a stepping stone to prove further results.

Second, since our model of adts is very close to star-free extended regular expressions (SEREs for short), that are star-free regular expressions extended with intersection and complementation, we provide a two-way translation from the former to the latter (Theorem 2). It is known that the class of languages denoted by SEREs coincides with the class of star-free languages [18], that can also be characterized as the class of languages definable in first-order logic over strings ( $\mathrm{FO}[<]$ ). We make explicit a translation from adts into $\mathrm{FO}[<]$ (Lemma 1) to shed light on the role played by the countermeasure-depth. Our translation is reminiscent of the constructions in [13] for an alternative proof of the result in [22] that relates the classic dot-depth hierarchy of star-free languages and the $\mathrm{FO}[<]$ quantifier alternation hierarchy. In particular, we show (Lemma 2) that any language definable by an adt with countermeasure-depth less than equal to $k$ is definable in $\Sigma_{k+1}$, the $(k+1)$-th level of the first-order quantifier alternation hierarchy.
Starting from the proof used in 22 to show the strictness of the dot-depth hierarchy, we demonstrate that there exists an infinite family of languages whose definability by an adt requires arbitrarily large countermeasure-depths. It should be noticed that our notion of countermeasure-depth slightly differs from the complementation-depth considered in [21] for extended regular expressions ${ }^{4}$, because the new operator AnD is rather a relative complementation. As a result, the countermeasure-depth of adts induces a new hierarchy of all star-free languages, that we call the ADT-hierarchy, that coincides (at least on the very first levels) neither with the dot-depth hierarchy, nor with the first-order logic quantifier alternation hierarchy.

Third, we study three natural decision problems for adts, namely the membership problem ( $A D T$-MEMB), the non-emptiness problem ( $A D T$-NE) and the equivalence problem ( $A D T$-EQUIV). The problem $A D T$-MEMB is to determine if

[^1]a trace is in the semantics of an adt. From a practical security point of view, $A D T$-memb addresses the ability to recognize an attack, say, in a log file. The problem $A D T$-NE consists of, given an adt, deciding if its semantics is nonempty. Otherwise said, whether the information system can be attacked or not. Finally, the problem $A D T$-EQUIV consists of deciding whether two adts describe the same attacks or not. Our results are summarized in Table 1 .

The paper is organized as follows. Section 2 proposes an introductory example. Next, we define our model of adts in Section 3, their trace semantics and countermeasure-depth, and present the Small Model Property for adts with countermeasure-depth bounded by one. We then show in Section 4 that adts coincide with star-free languages. We next study the novel hierarchy induced by the countermeasure-depth (Section 5) and study decision problems on adts (Section 6).

## 2 Introductory Example

Consider a thief (the proponent) who wants to steal two documents inside two different safes (Safe 1 and Safe 2), without being seen. The safes are located in two different but adjacent rooms (Room 1 and Room 2) in a building; the entrance/exit door of the building leads to Room 1. The rooms are separated by a door and each room has a window. Initially, the thief is outside of the building. A strategy for the proponent to steal the documents is to attempt to open Safe 1 until it succeeds, then open Safe 2 until it succeeds (and finally to exit the building). However, this strategy can be easily countered by the company, say by hiring a security guard visiting the rooms on some regular basis.

Security experts would commonly use an adt to describe how the proponent may achieve her goal and, at the same time, the ways its opponent (the company) may prevent the proponent from reaching her goal. An informal adt expressing the situation is given in Figure 1a, where traditionally goals of the proponent are represented in red circles, while countermeasures of the opponent are represented in green squares. An arrow from a left sibling to a right sibling specifies that the former goal must be achieved before starting the latter. A countermeasure targeting a proponent goal is represented with dashed lines.

As said, the graphical model of Figure 1a is informal and cannot be exploited by any automated tool for reasoning. With the setting proposed in this contribution, we make it formal, and in particular we work out a new binary operator CO for "countermeasure", graphically reflected with a curved dashed line between two siblings: the left sibling is a proponent's goal while the right sibling is the opponent's countermeasure - with this convention, we can unambiguously retrieve the player's type of an adt node. The more formal version of the adt in Figure 1a is drawn in Figure 1b (details for its construction can be found in Example 4).

The proposed semantics for adts also allows us to consider nested co operators to express countermeasures of countermeasures. For example, a proponent countermeasure against the company countermeasure could be to be disguised


Fig. 1: Adts for the thief and company problem.
as an employee working in the building; we formalise this situation in Example 4 It should be observed that nested countermeasure is a core aspect of our contribution and the main subject of Section 5 .

## 3 Attack-Defense Trees and Countermeasure-Depth

Preliminary notations For the rest of this paper, we fix Prop a finite set of propositions and we assume that the reader is familiar with propositional logic. We use typical symbol $\gamma$ for propositional formulas over Prop and write a valuation of the propositional variables $a$ as an element of $\Sigma:=2^{\text {Prop }}$, that will be viewed as an alphabet. A trace $t$ over Prop, is finite word over $\Sigma$, that is a finite sequence of valuations. We denote the empty trace by $\varepsilon$, and we define $\Sigma^{+}:=\Sigma^{*} \backslash\{\varepsilon\}$. For a trace $t=a_{1} \ldots a_{n}$, we define $|t|:=n$, its the length, as the number of valuations appearing in $t$, and we let $t[i]:=a_{i}$, for each $i \in\{1, \ldots, n\}$. We define the classic concatenation of traces: given two traces $t=a_{1} \ldots a_{n}$ and $t^{\prime}=a_{1}^{\prime} \ldots a_{m}^{\prime}$, we define $t \cdot t^{\prime}:=a_{1} \ldots a_{n} a_{1}^{\prime} \ldots a_{m}^{\prime}$. We also lift this operator to sets of traces in the usual way: given two sets of traces $L$ and $L^{\prime} \subseteq \Sigma^{*}$, we let $L \cdot L^{\prime}:=\left\{t \cdot t^{\prime}: t \in L\right.$ and $\left.t^{\prime} \in L^{\prime}\right\}$. For a trace $t=a_{1} \ldots a_{n} \in \Sigma^{*}$ and $1 \leq i \leq n$, the trace $t^{\prime}=a_{1} \ldots a_{i}$ is a prefix of $t$, written $t^{\prime} \preceq t$.

We define attack-defense trees (adts) over Prop, as well as their trace semantics and their countermeasure-depth, and develop enlightening examples. Adts are standard labeled finite trees with a dedicated set of labels based on the special $\epsilon$ label and propositional formulas for leaves and on the set \{OR, SAND, AND, CO \} for internal nodes.

Definition 1. The set $A D T$ of adts over Prop is inductively defined by:

- the empty-word leaf $\epsilon$ and every propositional formula $\gamma$ over Prop are in ADT;
- if trees $\tau_{1}, \ldots, \tau_{n}$ are in $A D T$, so are $\operatorname{OR}\left(\tau_{1}, \ldots, \tau_{n}\right), \operatorname{SAND}\left(\tau_{1}, \ldots, \tau_{n}\right)$, and $\operatorname{AND}\left(\tau_{1}, \ldots, \tau_{n}\right)$;
- if trees $\tau$ and $\tau^{\prime}$ are in $A D T$, so is $\operatorname{CO}\left(\tau, \tau^{\prime}\right)$.

The size of an adt $\tau$, written $|\tau|$ is defined as the sum of the sizes of its leaves, provided the size of $\epsilon$ is 1 , while the size of $\gamma$ is its size when seen as a propositional formula.

Regarding the semantics, adts describe a set of traces over alphabet $\Sigma:=$ $2^{\text {Prop }}$, hence their trace semantics. Formally, for an adt $\tau$, we define the language $\operatorname{Traces}(\tau) \subseteq \Sigma^{*}$. First, we set $\operatorname{Traces}(\epsilon)=\{\epsilon\}$. Now, a leaf adt hosting formula $\gamma$ denotes the reachability goal $\gamma$, that is the set of traces ending in a valuation satisfying $\gamma$ (we use the classic notations $T=p \vee \neg p$ and $\perp=\neg \top$ with $p \in$ Prop). We make our trace semantics compositional by providing the semantics of the four operators OR, SAND, CO, and AND in terms of how the subgoals described by their arguments interact. Operator OR tells that at least one of the subgoals has to be achieved. Operator SAND requires that all the subgoals need being achieved in the left-to-right order. The binary operator co requires to achieve the first subgoal without achieving the second one. Finally, operator AND tells that all subgoals need to be achieved, regardless of the order. Without any countermeasure, AND can be seen as a relaxation of the SAND, but it is not true in general (see example 3).

At the level of the property described by an adt, i.e. a trace language, the operators correspond to specific language operations: OR corresponds to union, SAND to concatenation, CO to a relativized complementation. Only AND corresponds to a less classic operation: a trace $t$ belongs to the language of $\operatorname{AND}\left(\tau_{1}, \tau_{2}\right)$ if $t$ belongs to the language of $\tau_{1}$ and has a prefix in the language of $\tau_{2}$ or viceversa. Formally:

Definition 2. Let $L_{1}, L_{2}$ be two languages over alphabet $\Sigma$. The each of $L_{1}$ and $L_{2}$ is the language $L_{1} \sqcap L_{2}:=\left(L_{1} \cdot \Sigma^{*} \cap L_{2}\right) \cup\left(L_{2} \cdot \Sigma^{*} \cap L_{1}\right)$.

Because operator $\sqcap$ is associative, we can define $L_{1} \sqcap L_{2} \sqcap \ldots \sqcap L_{n}$ that amounts to being equal to $\bigcup_{i \in\{1, \ldots, n\}}\left(L_{i} \cap \bigcap_{j \neq i} L_{j} \cdot \Sigma^{*}\right)$.
Example 1. A word $w=a_{1} a_{2} \ldots a_{m}$ belongs to $L_{1} \sqcap L_{2} \sqcap L_{3}$ whenever there are three (possibly equal) positions $i_{1}, i_{2}, i_{3}=m$ such that, for each $j \in\{1,2,3\}$, the word prefix $a_{1} \ldots a_{i_{j}} \in L_{\pi(j)}$, for some permutation $\pi$ of $(1,2,3)$.

We can now formally define the adt semantics.

## Definition 3.

$-\operatorname{Traces}(\epsilon):=\{\varepsilon\}$ and $\operatorname{Traces}(\gamma):=\left\{a_{1} \ldots a_{n} \in \Sigma^{*}: a_{n} \models \gamma\right\} ;$ In particular, $\operatorname{Traces}(\top)=\Sigma^{+}$and $\operatorname{Traces}(\perp)=\emptyset$;
$-\operatorname{Traces}\left(\operatorname{OR}\left(\tau_{1}, \ldots, \tau_{n}\right)\right):=\operatorname{Traces}\left(\tau_{1}\right) \cup \ldots \cup \operatorname{Traces}\left(\tau_{n}\right)$;
$-\operatorname{Traces}\left(\operatorname{SanD}\left(\tau_{1}, \ldots, \tau_{n}\right)\right):=\operatorname{Traces}\left(\tau_{1}\right) \cdot \ldots \cdot \operatorname{Traces}\left(\tau_{n}\right) ;$
$-\operatorname{Traces}\left(\mathrm{Co}\left(\tau_{1}, \tau_{2}\right)\right):=\operatorname{Traces}\left(\tau_{1}\right) \backslash \operatorname{Traces}\left(\tau_{2}\right) ;$

$$
-\operatorname{Traces}\left(\operatorname{AND}\left(\tau_{1}, \ldots, \tau_{n}\right)\right)=\operatorname{Traces}\left(\tau_{1}\right) \sqcap \ldots \sqcap \operatorname{Traces}\left(\tau_{n}\right)
$$

In the rest of the paper, we say for short that an adt is non-empty, written $\tau \neq \emptyset$, whenever $\operatorname{Traces}(\tau) \neq \emptyset$. We say that two adts $\tau$ and $\tau^{\prime}$ are equivalent, whenever $\operatorname{Traces}(\tau)=\operatorname{Traces}\left(\tau^{\prime}\right)$.

Remark 1. Since all operators $\cup, \cdot$ and $\sqcap$ over trace languages are associative, the trees of the form $\mathrm{OP}\left(\tau_{1}, \mathrm{OP}\left(\tau_{2}, \tau_{3}\right)\right), \mathrm{OP}\left(\mathrm{OP}\left(\tau_{1}, \tau_{2}\right), \tau_{3}\right)$, and $\mathrm{OP}\left(\tau_{1}, \tau_{2}, \tau_{3}\right)$ are all equivalent, when $O P$ ranges over \{OR, SAND, AND \}. As a consequence, we may sometime assume that nodes with such operators are binary.

We now introduce some notations for particular adts to ease our exposition and provide some examples of adts with their corresponding trace property.

We define a family of adts of the form $|\geq \ell|,|<\ell|,|=\ell|$, where $\ell$ is a nonzero natural. We let $|\geq \ell|:=\operatorname{SAND}(T, \ldots, T)$ where $T$ occurs $\ell$ times; $|<\ell|:=$ $\operatorname{CO}(T,|\geq \ell|)$; and $|=\ell|:=\operatorname{CO}(|\geq \ell|,|\geq \ell+1|)$. It is easy to establish that adt $|\geq \ell|$ (resp. $|<\ell|,|=\ell|$ ) denotes the set of traces of length at least (resp. at most, exactly) $\ell$. We also consider particular adts and constructs for them.

- ALL $:=\mathrm{OR}(\epsilon, \top)$,
$-\operatorname{Not}(\tau):=\operatorname{Co}(\operatorname{ALL}, \tau)$, and $\operatorname{INTER}\left(\tau_{1}, \tau_{2}\right):=\operatorname{NOT}\left(\operatorname{OR}\left(\operatorname{NOT}\left(\tau_{1}\right), \operatorname{NOT}\left(\tau_{2}\right)\right)\right)$,
$-\rceil \tau\lceil:=\operatorname{SAND}(\mathrm{ALL}, \tau, \operatorname{ALL}),\rceil \tau:=\operatorname{SAND}(\mathrm{ALL}, \tau)$ and $\tau\lceil:=\operatorname{SAND}(\tau, \operatorname{ALL})$.
- Given a formula $\gamma$ over Prop, we let $\underline{\gamma}:=\operatorname{CO}(\gamma,|\geq 2|)$.

Based on these notations, we develop further examples.
Example 2. $\quad-\operatorname{Traces}(\mathrm{ALL})=\Sigma^{*} ;$
$-\operatorname{Traces}(\operatorname{Not}(\tau))=\Sigma^{*} \backslash \operatorname{Traces}(\tau) ;$
$-\operatorname{Traces}\left(\operatorname{INTER}\left(\tau_{1}, \tau_{2}\right)\right)=\operatorname{Traces}\left(\tau_{1}\right) \cap \operatorname{Traces}\left(\tau_{2}\right)$;

- Traces $(\underline{\gamma})$ is set of one-length traces whose unique valuation satisfies $\gamma$; in particular, when a valuation $a$ is understood as a formula, namely formula $\bigwedge_{p \in a} p \wedge \bigwedge_{p \notin a} \neg p$, the adt $\underline{a}:=\operatorname{CO}(a,|\geq 2|)$ is such that $\operatorname{Traces}(\underline{a})=\{a\} ;$
- For a valuation $a$,
- Traces $\left(7 \underline{a}\left)=\operatorname{Traces}( \rceil a\lceil )=\operatorname{Traces}\left(a\lceil )=\Sigma^{*} a \Sigma^{*} ;\right.\right.\right.$
- Traces $( \rceil \underline{a})=\operatorname{Traces}( \rceil a)=\operatorname{Traces}(a)=\Sigma^{*} a$;
- also, $\operatorname{Traces}\left(\underline{a}\left)=a \Sigma^{*}\right.\right.$.

Example 3. Note that AND cannot be seen as a kind of relaxation of SAND. For the set of propositions $\{p\}$, if we consider the formula $p$ as a leaf, $\operatorname{Traces}(\underline{p})=\{p\}$ and $\operatorname{Traces}(\neg p)=\{\emptyset\}$. Thus $\operatorname{SanD}(\neg p, \underline{p})=\{t\}$ with trace $t=\emptyset p$. However $\operatorname{AND}(\neg p, \underline{p})=\bar{\emptyset}$. Let us notice that the construction of this example uses the co operator (hidden in $\underline{p}$ and $\neg \underline{p}$ ).

Example 4. We come back to the situation of our introductory example (Section 21. First, we discuss the formal semantics of the informal tree in fig. 1a. To do so, we propose the following set of propositions: $\operatorname{Prop}=\left\{E, S_{1}, S_{2}, G\right\}$ where $E$ holds when the thief is entering the building, $S_{1}$ (resp. $S_{2}$ ) holds when the first
(resp. second) safe is open, and $G$ is true if a guard is in the building. The situation can be described by the following adt: $\tau_{e x_{1}}=\operatorname{SAND}\left(E, \operatorname{Co}\left(S_{1} \wedge S_{2}, G\lceil )\right)\right.$, represented in Figure 1b, where we distinguish SAND with a curved line and CO with a dashed line. We have $\operatorname{Traces}\left(\tau_{e x_{1}}\right)=\left\{a_{1} \ldots a_{n} \in \Sigma^{*}: a_{n} \models E\right\} \cdot\left(\left\{a_{1} \ldots a_{n} \in\right.\right.$ $\left.\Sigma^{*}: a_{n} \models S_{1} \wedge S_{2}\right\} \backslash\left\{a_{1} \ldots a_{n} \in \Sigma^{*}: \exists i\right.$ such that $\left.a_{i} \models G\right\}$ ). If we write $\gamma_{\varphi}=\left\{a \in 2^{\text {Prop }}: a \models \varphi\right\}$, we have $\operatorname{Traces}\left(\tau_{e x_{1}}\right)=\Sigma^{*} \gamma_{E} \cdot\left(\Sigma^{*} \gamma_{S_{1} \wedge S_{2}} \backslash \Sigma^{*} \gamma_{G} \Sigma^{*}\right)$. In other words, we want all traces where $E$ holds at some point and, after it, $G$ cannot be true and finish by a valuation where $S_{1} \wedge S_{2}$ holds.

In order to illustrate the nesting of countermeasures, we now allow the thief to disguise himself as an employee (assuming that when disguised, the guard does not identify him as a thief). To do so, we extended the set of propositions: Prop $^{\prime}=\left\{E, S_{1}, S_{2}, G, D\right\}$, where $D$ holds when the thief is disguised. The situation is now described by the following adt: $\tau_{\text {ex }}=\operatorname{SAND}\left(E, \operatorname{Co}\left(S_{1} \wedge\right.\right.$ $\left.S_{2},\right\rceil \mathrm{Co}(G, D)))$. The semantics for $\tau_{e x_{2}}$ is all traces where $E$ holds at some point and, after it, $G$ cannot be true, except if $D$ holds at the same time, and finish by a valuation where $S_{1} \wedge S_{2}$ holds. A representation of $\tau_{e x_{2}}$ can be found in Appendix A

We now stratify the set ADT of adts according to their countermeasure-depth that denotes the maximum number of nested countermeasures.

Definition 4. The countermeasure-depth of an adt $\tau$, written $\delta(\tau)$, is inductively defined by:
$-\delta(\epsilon):=\delta(\gamma)=0$;
$-\delta\left(O P\left(\tau_{1}, \ldots, \tau_{n}\right)\right):=\max \left\{\delta\left(\tau_{1}\right), \ldots, \delta\left(\tau_{n}\right)\right\}$ for every $O P \in\{\mathrm{OR}, \mathrm{SAND}, \mathrm{AND}\} ;$
$-\delta\left(\operatorname{Co}\left(\tau_{1}, \tau_{2}\right)\right):=\max \left\{\delta\left(\tau_{1}\right), \delta\left(\tau_{2}\right)+1\right\}$
We let $A D T_{k}:=\{\tau \in A D T: \delta(\tau) \leq k\}$ be the set of adts with countermeasuredepth at most $k$. Clearly $A D T_{0} \subseteq A D T_{1} \subseteq \ldots A D T_{k} \subseteq A D T_{k+1} \subseteq \ldots$, and $A D T=\bigcup_{k \in \mathbb{N}} A D T_{k}$.

Example 5. We list a couple of examples. $\delta(\operatorname{NOT}(\tau))=1+\delta(\tau) ; \delta\left(\operatorname{INTER}\left(\tau_{1}, \tau_{2}\right)\right)=$ $2+\max \left\{\delta\left(\tau_{1}\right), \delta\left(\tau_{2}\right)\right\} ; \delta\left(\operatorname{co}\left(\operatorname{co}\left(\gamma_{1}, \gamma_{2}\right), \gamma_{3}\right)\right)=1$ while $\delta\left(\operatorname{Co}\left(\gamma_{3}, \operatorname{Co}\left(\gamma_{2}, \gamma_{3}\right)\right)\right)=2$; $\delta(|<\ell|)=\delta(|=\ell|)=1$ while $\delta(|\geq \ell|)=0 ; \delta(\underline{\gamma})=1$; In Example 4. $\delta\left(\tau_{e x_{1}}\right)=1$ and $\delta\left(\tau_{e x_{2}}\right)=2$.

Also, $\delta(|\geq \ell|)=0, \delta(|<\ell|)=\delta(|=\ell|)=1$, so that $|\geq \ell| \in A D T_{0} ;|<\ell|$ and $|=\ell| \in A D T_{1}$. Moreover, $\tau_{e x_{1}} \in A D T_{1}$ and $\tau_{e x_{2}} \in A D T_{2}$.

We say that a language $L$ is $A D T_{k}$-definable (resp. $A D T$-definable), written $L \in A D T_{k}$ (resp. $A D T$ ), whenever $\operatorname{Traces}(\tau)=L$, for some $\tau \in A D T_{k}$ (resp. for some $k$ ).

It can be established that non-empty adts in $A D T_{1}$ enjoy small traces, i.e. smaller than the size of the tree.

Theorem 1 (Small model property for $A D T_{1}$ ). An adt $\tau \in A D T_{1}$ is non-empty if, and only if, there is a trace $t \in \operatorname{Traces}(\tau)$ with $|t| \leq|\tau|$.

We here only sketch the proof, whose details can be found in Appendix B. The technique we employ consists in defining a slight variant of the classic relation of super-word in language theory, that we call the lift binary relation. We prove that if $\tau \in A D T_{1}$, then there exists a finite set of generators, denoted $\operatorname{gen}(\tau)$, which is sufficient to describe $\operatorname{Traces}(\tau)$ through the lift relation. Next, we can prove that the traces in $\operatorname{gen}(\tau)$ have size bounded by the number of leaves of $\tau$. Notice that the result also holds for $A D T_{0}$, since $A D T_{0} \subseteq A D T_{1}$.

## 4 Adts, Star-free Languages, and First-Order Logic

We prove that adts coincide with star-free languages and first order formulas.

### 4.1 Reminders on Star-free Languages and First-Order Logic

The class of star-free languages introduced by 11|415 (over alphabet $\Sigma$ ) is obtained from the finite languages (or alternatively languages consisting of a single one-length word in $\Sigma$ ) by finitely many applications of Boolean operations ( $\cup$, $\cap$ and $\sim$ for the complement) and the concatenation product (see [18, Chapter 7]). Alternatively, one characterizes star-free languages by first considering extended regular expressions - that are regular expressions augmented with intersection and complementation, and second by restricting to star-free extended regular expressions (SEREs, for short) that are extended regular expressions with no Kleene-star operator. Regarding computational complexity aspects, we recall the following the subclass of SERES of extended regular expressions. The word membership problem (i.e., whether a given word belongs to the language denoted by a SERE) is in Ptime [9, Theorem 2], while the non-emptiness problem (i.e., is the denoted language empty?) and the equivalence problem (i.e., do two SEREs denote the same language?) are hard, both non-elementary [21, p. 162].

We now recall classical results on the first-order logic on finite words $\mathrm{FO}[<]$ (see details in [14, Chapter 29]). The signature of $\mathrm{FO}[<]$, say for words over an alphabet $\Sigma$, is composed of a unary predicate $a(x)$ for each $a \in \Sigma$, whose meaning is the "letter at position $x$ of the word is $a$ ", and the binary predicate $x<y$ that states "position $x$ is strictly before position $y$ in the word". For a $\mathrm{FO}[<]$-formula, we define its size $|\psi|$ as the size of the expression considered as a word. A language $L$ is $\mathrm{FO}[<]$-definable whenever there exists a $\mathrm{FO}[<]$-formula $\psi$ such that a word $w \in L$ if, and only if, $w$ is a model of $\psi$. Similarly, we say that an adt $\tau$ is $\mathrm{FO}[<]$-definable if $\operatorname{Traces}(\tau)$ is $\mathrm{FO}[<]$-definable. It is well-known that $\mathrm{FO}[<]$-definable languages coincide with star-free languages [11|22|13].

Also, for a fine-grained inspection of $\mathrm{FO}[<]$, let us denote by $\Sigma_{\ell}$ (resp. $\Pi_{\ell}$ ) the fragments of $\mathrm{FO}[<]$ consisting of formulas with at most $\ell$ alternation of $\exists$ and $\forall$ quantifier blocks, starting with $\exists$ (resp. $\forall$ ). The folklore results regarding satisfiability of $\mathrm{FO}[<]$-formulas [21|12] are: (a) The satisfiability problem for $\mathrm{FO}[<]$ is non-elementary; (b) The satisfiability for $\Sigma_{\ell}$ is in $(\ell-1)$-ExpsPACE ${ }_{5}^{5}$

[^2]We are not aware of any result that establishes a tight lower bound complexity for the satisfiability problem on the $\Sigma_{\ell}$ fragments of $\mathrm{FO}[<]$.

We lastly recall the definition of the dot-depth hierarchy of star-free languages: level 0 of this hierarchy is $B_{0}:=\left\{L \subseteq 2^{\Sigma}: L\right.$ is finite or co-finite $\}$, and evel $\ell$ is $B_{\ell}:=\left\{L \subseteq 2^{\Sigma}: L\right.$ is a Boolean combination of languages of the form $L_{1} \cdot \ldots \cdot L_{n}$ where $\left.L_{1}, \ldots, L_{n} \in B_{\ell-1}\right\}$. The dot-depth hierarchy has a tight connection with $\mathrm{FO}[<]$ fragments [22]: for every $\ell>0, \Sigma_{\ell} \subseteq B_{\ell} \subseteq \Sigma_{\ell+1}$.

### 4.2 Expressiveness of Adts

The first result of this section consist in showing that adts and star-free extended regular expressions share the same expressiveness.

Theorem 2. A language $L$ is star-free if, and only if, $L$ is ADT-definable.
For the "only if" direction of Theorem 2, we reason by induction on the class of star-free languages. For a language of the form $\{a\}$ where $a \in \Sigma$ one can take the adt $\underline{a}$ (that is $\operatorname{CO}(a,|\geq 2|)$ ). Now we can inductively build adequate adts for compound star-free languages by noticing that language operations of union and concatenation are captured by adts operators OR and SAND respectively, while complementation and intersection are obtained from the NOT(.) and INTER(.,.) as formalized in Example 2. One easily verifies that that the size of the adt corresponding to an SERE $E$ is in $O(|E|$ ), where $|E|$ denotes the size (number of characters) of $E$.

For the "if" direction of Theorem 2, it is easy to translate an adt into an SERE: the leaf $\epsilon$ translates into $\epsilon$, a leaf adt $\gamma$ translates into $\bigcup_{a \models \gamma} a$ - notice that this translation is exponential. For non-leaf adts, since every operator occurring in the adt has its language-theoretic counterpart the translation goes smoothly. However, the translation is exponential because of the adt operator AND, see Definition 2 .

We now dig into the ADT-hierarchy induced by the countermeasure-depth and compare it with the $\mathrm{FO}[<]$ fragments $\Sigma_{\ell}$ and $\Pi_{\ell}$.

We first step design a translation from ADT into $\mathrm{FO}[<]$, inductively over adts. The translation of an adt $\tau$ is written $\psi_{\tau}$. For the base cases of adts $\epsilon$ and $\gamma$, and we let: $\psi_{\epsilon}:=\forall x \perp$ and $\psi_{\gamma}:=\exists x\left(\forall y \neg(x<y) \wedge \bigvee_{a \models \gamma} a(x)\right)$.

Now, regarding compound adts, and not surprisingly, operator OR is reflected by the logical disjunction: $\psi_{\mathrm{OR}\left(\tau_{1}, \tau_{2}\right)}:=\psi_{\tau_{1}} \vee \psi_{\tau_{2}}$, while operator CO is reflected by means of the logical conjunction with the negated second argument: $\psi_{\mathrm{Co}\left(\tau_{1}, \tau_{2}\right)}:=$ $\psi_{\tau_{1}} \wedge \neg \psi_{\tau_{2}}$. On the contrary, the two remaining operators SAND and AND require to split the trace into pieces, which can be captured by the folklore operation of left (resp. right) position relativizations of $\mathrm{FO}[<]$-formulas w.r.t. a position [13, Proposition 2.1] (see also formulas of the form $\phi^{[x, y]}$ in [19]). Formally, given a position $x$ in the trace $t$ and an $\mathrm{FO}[<]$-formula $\psi$, we define formula $\psi \leq x$ (resp. $\psi^{>x}$ ) that holds of $t$ if the prefix (resp. suffix) of $t$ up to (resp. from) position $x$ satisfies $\psi$, as follows. For $\bowtie \in\{\leq,>\}$, we let:

$$
\begin{aligned}
& a(y)^{\bowtie x}=a(y) \\
& (y<z)^{\bowtie x}=(y<z) \\
& p(y)^{\bowtie x}=p(y)
\end{aligned}
$$

$$
\begin{aligned}
& \left(\psi \vee \psi^{\prime}\right)^{\bowtie x}=\psi^{\bowtie x} \vee \psi^{\prime \bowtie x} \\
& (\neg \psi)^{\bowtie x}=\neg \psi^{\bowtie x} \\
& (\exists y \psi)^{\bowtie x}=\exists y\left(y \bowtie x \wedge \psi^{\bowtie x}\right)
\end{aligned}
$$

Additionally, we write $\psi^{\leq 0}$ and $\psi^{>0}$ as the formulas obtained from $\psi^{\leq x}$ and $\psi^{>x}$ by replacing every occurrence of expressions $y \leq x$ and $y>x$ by $\perp$ and $\top$ respectively.

Remark 2. For every formula $\psi \in \Sigma_{\ell}$, we also have $\psi^{\leq x}, \psi^{>x} \in \Sigma_{\ell}$.
We can now complete the translation from ADT into $\mathrm{FO}[<]$ by letting (w.l.o.g., by Remark 1, we can consider binary SAND and AND):

$$
\begin{aligned}
& \psi_{\mathrm{SAND}\left(\tau_{1}, \tau_{2}\right)}:=\exists x\left[\psi_{\tau_{1}} \leq x \wedge \psi_{\tau_{2}}{ }^{>x}\right] \vee\left(\psi_{\tau_{1}} \leq 0 \wedge \psi_{\tau_{2}}\right) \\
& \psi_{\mathrm{AND}\left(\tau_{1}, \tau_{2}\right)}:=\exists x\left[\left(\psi_{\tau_{1}} \leq x \wedge \psi_{\tau_{2}}\right) \vee\left(\psi_{\tau_{2}} \leq x \wedge \psi_{\tau_{1}}\right)\right] \vee\left(\psi_{\tau_{1}} \leq 0 \wedge \psi_{\tau_{2}}\right) \vee\left(\psi_{\tau_{2}} \leq 0 \wedge \psi_{\tau_{1}}\right) .
\end{aligned}
$$

Lemma 1. - For any $t \in \Sigma^{*}, t \equiv \psi_{\tau}$ iff $t \in \operatorname{Traces}(\tau)$.

- For any adt $\tau$, formula $\psi_{\tau}$ is of size exponential in $|\tau|$.

In the rest of the paper, we use mere inclusion symbol $\subseteq$ between subclasses of ADT and subclasses of $\mathrm{FO}[<]$, with the canonical meaning regarding the denoted trace languages.

An accurate inspection of the translation $\tau \mapsto \psi_{\tau}$ entails that every $A D T_{k^{-}}$ definable adt can be equivalently represented by a $\Sigma_{k+2}$-formula, namely $A D T_{k} \subseteq$ $\Sigma_{k+2}$. However, we significantly refine this expressiveness upperbound for $A D T_{k}$.

Lemma 2. 1. $A D T_{0} \subseteq \Sigma_{2} \cap \Pi_{2}$ - with an effective translation.
2. For every $k>0, A D T_{k} \subseteq \Sigma_{k+1}$ - with an effective translation.

Regarding Item 1 of Lemma 2, it can be observed that, whenever $\tau \in A D T_{0}$, the quantifiers $\forall$ and $\exists$ commute in $\psi_{\tau}$ (see Appendix C). Now, for Item 2, the proof is conducted by induction over $k$ (see Appendix C). We sketch here the case $k>1$. First, remark that if $\tau_{1}, \ldots, \tau_{n} \in A D T_{k-1}$ are $\Sigma_{k}$-definable, then $\operatorname{OR}\left(\tau_{1}, \ldots, \tau_{n}\right), \operatorname{SAND}\left(\tau_{1}, \ldots, \tau_{n}\right)$ and $\operatorname{AND}\left(\tau_{1}, \ldots, \tau_{n}\right)$ remain $\Sigma_{k}$-definable, as formulas are obtained from conjunctions or disjunctions of $\Sigma_{k}$-definable formulas. Moreover, if $\tau_{1}$ and $\tau_{2}$ are $\Sigma_{k}$-definable, then $\tau=\operatorname{CO}\left(\tau_{1}, \tau_{2}\right)$ is $\Sigma_{k+1}$-definable as a formula can be obtained from a boolan combination of two $\Sigma_{k}$-definable formulas. Finally, with $k$ still fixed, it can be shown by induction over the size of an adt $\tau$ that, if $\tau \in A D T_{k}$, since all its countermeasures operators are of the form $\operatorname{CO}\left(\tau_{1}, \tau_{2}\right)$ where $\tau_{1} \in A D T_{k}$ and $\tau_{2} \in A D T_{k-1}$, we have that adt $\tau$ is also $\Sigma_{k+1}$-definable, which concludes.

We can also establish lowerbounds in the ADT-hierarchy.
Lemma 3. 1. $B_{\ell} \subseteq A D T_{2 \ell+2}$, and therefore $\Sigma_{\ell} \subseteq A D T_{2 \ell+2}$
2. $\Sigma_{1} \subseteq A D T_{0}$ - with an effective translation.

Item 1 of Lemma 3 is obtained by an induction of $\ell$. Regarding Item 2, the translation consists in putting the main quantifier-free subformula of a $\Sigma_{1}$ formula in disjunctive normal form, and to focus for each conjunct on the set of "ordering" literals of the form $x<y$ or $\neg(x<y)$ (leaving aside the other literals of the form $a(x)$ or $\neg a(x)$ for a while). Each ordering literal naturally induces a partial order between the variables. We expend this partial order constraint over the variables as a disjunction of all its possible linearizations. For example, the conjunct $x<y \wedge \neg(y<z)$ is expended as the equivalent formula $(x<y \wedge y=$ $z) \vee(x<z \wedge z<y) \vee(z=x \wedge x<y) \vee(z<x \wedge x<y)$. Now, each disjunct of this new formula, together with the constraints $a(x)$ (or $\neg a(x)$ ), can easily be specified by a SAND-rooted adt (ie. the root is a SAND). The initial $\Sigma_{1}$-formula then is associated with the or-rooted tree that gathers all the aforementioned SAND-rooted subtrees (see Remark 3in Appendix B). Notice that the translation may induce at least an exponential blow-up.

The reciprocal of Item 2 in Lemma 3 is an open question. Still, we have little hope that it holds because $\psi_{\gamma}$ seems to require a property expressible in $\Sigma_{2} \backslash \Sigma_{1}$.

## 5 Strictness of the ADT-Hierarchy

One can notice that the adt $|<2| \in A D T_{1}$ defines the finite language of traces of length at most 1 , while it can be established that languages arising from adts in $A D T_{0}$ are necessarily infinite - if inhabited by a non-empty word (see Lemma 4 of Appendix B). Thus, we can easily deduce $A D T_{0} \subsetneq A D T_{1}$. This section aims at showing that the entire ADT-hierarchy is strict:

Proposition 1. For every $k \in \mathbb{N}, A D T_{k} \subsetneq A D T_{k+1}$, even if Prop $=\{p\}$.
To show that $A D T_{k} \subsetneq A D T_{k+1}$, we use a family of languages, originally introduced in [23], that we write $\left\{W_{k}\right\}_{k \in \mathbb{N}}$ over the two-letter alphabet $\Sigma$ obtained from Prop $=\{p\}$. For readability, we use symbol $a$ (resp. b) for the valuation $\{p\}$ (resp. $\emptyset$ ) of $\Sigma$. Formally, we define $W_{k} \subseteq \Sigma^{*}$ as follows - where $\|w\|$ denotes the number of occurrences of $a$ minus the number of occurrences of $b$ in the word $w$ :

We let $w \in W_{k}$ whenever all the following holds.
$-\|w\|=0 ;$

- for every $w^{\prime} \preceq w, 0 \leq\left\|w^{\prime}\right\| \leq k$;
- there exists $w^{\prime \prime} \preceq w$ s.t. $\left\|w^{\prime \prime}\right\|=k$.

In [23, Theorem 2.1], it is shown that $W_{k} \in B_{k} \backslash B_{k-1}$, for all $k \geq 1$.
We now determine the position of $W_{k}$ languages in the ADT-hierarchy. We show Proposition 2.

Proposition 2. For each $k>0, W_{k} \in A D T_{k+1} \backslash A D T_{k-2}$.
First, $W_{k} \notin A D T_{k-2}$ because $W_{k}$ is not $\Sigma_{k-1}$-definable ([23]). By an inductive argument over $k$ (see Appendix D), we can build an adt $\mu_{k} \in A D T_{k+1}$
that captures $W_{k}$. We only sketch here the case $k=1$. For $W_{1}$, we set $\mu_{1}:=$ $\operatorname{co}\left(b, \operatorname{OR}\left(\underline{b}\lceil,\rceil \operatorname{CO}(|\geq 2|, \operatorname{AND}(a\lceil, b\lceil )) \Gamma))\right.\right.$, depicted in Figure $2 a$ Note that $\delta\left(\mu_{1}\right)=$ 2 and that by a basic use of semantics, we have $\operatorname{Traces}\left(\mu_{1}\right)=(a b)^{+}=W_{1}$.


Now we have all the material to prove Proposition 1. Indeed, assuming the hierarchy collapses at some level will contradict Proposition 2. Notice that this argument is not constructive as we have no witness of $A D T_{k} \subsetneq A D T_{k+1}$ but for $k=1$ where it can be shown that $(a b)^{+} \in A D T_{2} \backslash A D T_{1}$ (see Appendix B, Example 7 and Proposition 3 ). Our results about the ADT-hierarchy are depicted on Figure 2b.

## 6 Decision Problems on Attack-Defense Trees

We study classical decision problems on languages, through the lens of adts, with a focus on the role played by the countermeasure-depth in their complexities. The problems are the following.

- The membership problem, written $A D T$-MEMB, is defined by:

Input: $\quad \tau$ an attack-defense tree and $t$ a trace.
Output: "YES" if $t \in \operatorname{Traces}(\tau)$, "NO" otherwise.

- The non-emptiness problem, written $A D T-\mathrm{NE}$, is defined by:

Input: $\tau$ an attack-defense tree.
Output: "YES" if $\operatorname{Traces}(\tau) \neq \emptyset$, "NO" otherwise.
We use notations $A D T_{k}$-MEMB and $A D T_{k}$-NE whenever the input adts of the respective decision problems are in $A D T_{k}$, with a fixed $k$. Our results are summarized in Table 1, and we below comment on them, row by row.

|  | $A D T_{0}$ | $A D T_{1}$ | $A D T_{k}(k \geq 2)$ | ADT |
| :---: | :---: | :---: | :---: | :---: |
| $A D T$-mEMB | Ptime | Ptime | Ptime | Ptime |
| $A D T$ - NE | NP-comp | NP-comp | $\begin{gathered} (k+1) \text {-EXPSPACE } \\ 1 \geq \operatorname{NSPACE}\left(g\left(k-5, c \sqrt{\frac{n-1}{3}}\right)\right) \end{gathered}$ | non-elem |
| $A D T$-EQUIV | coNP-comp | 4-Expspace | $\begin{gathered} (k+2) \text {-EXPSPACE } \\ 2 \geq \mathrm{NSPACE}\left(g\left(k-4, c \sqrt{\frac{n-1}{3}}\right)\right) \end{gathered}$ | non-elem |

${ }^{1}$ if $k \geq 5$.
${ }^{2}$ if $k \geq 4$.
Table 1: Computational complexities of decision problems on adts.

Regarding $A D T$-memb (first row of Table 1), we recall that adts and SEREs are expressively equivalent, but with a translation (Theorem 2) from the former to the latter that is not polynomial. We therefore cannot exploit [9, Theorem 2] for a PTIME complexity of the word membership problem for SEREs, and have instead developed a dedicated alternating logarithmic-space algorithm in Appendix E.

Regarding $A D T$-NE (second row of Table 1), and because SEREs can be translated as adts (see "if" direction in the proof of Theorem 2), the problem $A D T$-ne inherits from the hardness of the non-emptiness of SEREs 21, p. 162]. In its full generality, $A D T$-NE is therefore non-elementary (last column). Moreover, by our exponential translation of $A D T_{k}$ into the $\mathrm{FO}[<]$-fragment $\Sigma_{k+1}$ (Lemma 2), we obtain the ( $k+1$ )-ExPSPACE upper-bound complexity for $A D T_{k}-\mathrm{NE}$ (recall satisfiability problem for $\Sigma_{\ell}$ is $(\ell-1)$-EXPSPACE). Additionally, a lower-bound for $A D T_{k}$-NE is directly given by [21, Theorem 4.29]: for SERES that linearly translate as adts with countermeasure-depth $k$, their non-emptiness is at least $\operatorname{Nspace}\left(g\left(k-5, c \sqrt{\frac{n-1}{3}}\right)\right)$. Interestingly, the Small Model Property (Theorem1) yields an NP upper-bound complexity for $A D T_{1}-\mathrm{NE}$, also applicable for $A D T_{0}-\mathrm{NE}$, that is optimal since one can reduce the NP-complete satisfiability of propositional formulas to the non-emptiness of leaf adts.

Finally regarding $A D T$-EQUIV (last row of Table 11, one can observe that $A D T$-ne and $A D T$-EQUIV are very close. First, $A D T$-NE is a particular case of $A D T$-EQUIV with the second input adt $\tau_{2}:=\emptyset$. As a consequence, $A D T$-EQUIV inherits from the hardness of $A D T$-NE, and is therefore non-elementary (last column). Also, because deciding the equivalence between $\tau_{1}$ and $\tau_{2}$ amounts to
deciding whether $\operatorname{OR}\left(\operatorname{CO}\left(\tau_{1}, \tau_{2}\right), \operatorname{CO}\left(\tau_{2}, \tau_{1}\right)\right)=\emptyset$, we get reduction from $A D T_{k^{-}}$ EQUIV into $A D T_{k+1}$-NE which provides the results announced in Columns 1-3.

## 7 Discussion

First, we discuss our model of adts with regard to the literature. However, we do not compare with settings where adts leaves are actions [10], as they yield only finite languages, and address other issues [24].
Our adts have particular features, but remain somehow standard. Regarding the syntax, firstly, even though we did not type our nodes as proponent/opponent, the countermeasure operator fully determines the alternation between attack and defense. Also, we introduced the non-standard leaf $\epsilon$ for the singleton empty-trace language, not considered in the literature. Still, it is a very natural object in the formal language landscape, and anyhow does not impact our overall computational complexity analysis. Regarding the semantics, it can be shown that the definition of our operator AND together with the reachability goal semantics of the leaves coincides with the acknowledged semantics considered in 112 .

Second, we discuss our results. We showed the strictness of the $A D T$-hierarchy in a non-constructive manner. However, exhibiting an element of $A D T_{k+1} \backslash$ $A D T_{k}(k \geq 2)$ is still an open question. Since $W_{1} \in A D T_{2} \backslash A D T_{1}$, languages $W_{k}$ are natural candidates. We conjecture it is the case and we are currently working on Erenfeucht-Fraissé(EF)-like games for adts (in the spirit of [23] for SEREs) to prove it. Moreover, EF-like games for adts may also help to better compare the ADT-hierarchy and the $\mathrm{FO}[<]$ alternation hierarchy, in particular, whether the hierarchies eventually coincide. For now, finding a tighter inclusion of $\Sigma_{\ell}$ in some $A D T_{k}$ for each $\ell$ seems difficult; recall that we established $\Sigma_{\ell} \subseteq A D T_{2 \ell+2}$. Any progress in this line would be of great help to obtain tight complexity bounds for $A D T_{k}$-NE and $A D T_{k}$-EQUIV. Finally, determining the level of a language in the ADT-hierarchy seems as hard as determining its level in the dot-depth hierarchy, recognised as a difficult question [17|3].

## References

1. M. Audinot, S. Pinchinat, and B. Kordy. Is my attack tree correct? In European Symposium on Research in Computer Security, pages 83-102. Springer, 2017.
2. T. Brihaye, S. Pinchinat, and A. Terefenko. Adversarial formal semantics of attack trees and related problems. In P. Ganty and D. D. Monica, editors, Proceedings of the 13th International Symposium on Games, Automata, Logics and Formal Verification, GandALF 2022, Madrid, Spain, September 21-23, 2022, volume 370 of EPTCS, pages 162-177, 2022.
3. V. Diekert and P. Gastin. First-order definable languages. In J. Flum, E. Grädel, and T. Wilke, editors, Logic and Automata: History and Perspectives [in Honor of Wolfgang Thomas], volume 2 of Texts in Logic and Games, pages 261-306. Amsterdam University Press, 2008.
4. S. Eilenberg. Automata, languages, and machines. Academic press, 1974.
5. O. Gadyatskaya, R. R. Hansen, K. G. Larsen, A. Legay, M. C. Olesen, and D. B. Poulsen. Modelling attack-defense trees using timed automata. In Formal Modeling and Analysis of Timed Systems: 14th International Conference, FORMATS 2016, Quebec, QC, Canada, August 24-26, 2016, Proceedings 14, pages 35-50. Springer, 2016.
6. R. Jhawar, B. Kordy, S. Mauw, S. Radomirović, and R. Trujillo-Rasua. Attack trees with sequential conjunction. In IFIP International Information Security and Privacy Conference, pages 339-353. Springer, 2015.
7. B. Kordy, S. Mauw, S. Radomirović, and P. Schweitzer. Attack-defense trees. Journal of Logic and Computation, 24(1):55-87, 2014.
8. B. Kordy, M. Pouly, and P. Schweitzer. Computational aspects of attack-defense trees. In Security and Intelligent Information Systems: International Joint Conferences, SIIS 2011, Warsaw, Poland, June 13-14, 2011, Revised Selected Papers, pages 103-116. Springer, 2012.
9. O. Kupferman and S. Zuhovitzky. An improved algorithm for the membership problem for extended regular expressions. In Mathematical Foundations of Computer Science 2002: 27th International Symposium, MFCS 2002 Warsaw, Poland, August 26-30, 2002 Proceedings 27, pages 446-458. Springer, 2002.
10. S. Mauw and M. Oostdijk. Foundations of attack trees. In International Conference on Information Security and Cryptology, pages 186-198. Springer, 2005.
11. R. McNaughton and S. A. Papert. Counter-Free Automata (MIT research monograph no. 65). The MIT Press, 1971.
12. A. R. Meyer. Weak monadic second order theory of succesor is not elementaryrecursive. In Logic Colloquium: Symposium on Logic Held at Boston, 1972-73, pages 132-154. Springer, 2006.
13. D. Perrin and J.-E. Pin. First-order logic and star-free sets. Journal of Computer and System Sciences, 32(3):393-406, 1986.
14. J. Pin, editor. Handbook of Automata Theory. European Mathematical Society Publishing House, Zürich, Switzerland, 2021.
15. J. E. Pin and M. P. Schützenberger. Variétés de langages formels, volume 17. Masson Paris, 1984.
16. S. Pinchinat, M. Acher, and D. Vojtisek. Towards synthesis of attack trees for supporting computer-aided risk analysis. In C. Canal and A. Idani, editors, Software Engineering and Formal Methods - SEFM 2014 Collocated Workshops: HOFM, SAFOME, OpenCert, MoKMaSD, WS-FMDS, Grenoble, France, September 1-2, 2014, Revised Selected Papers, volume 8938 of Lecture Notes in Computer Science, pages 363-375. Springer, 2014.
17. T. Place and M. Zeitoun. The tale of the quantifier alternation hierarchy of firstorder logic over words. ACM SIGLOG News, 2(3):4-17, 2015.
18. G. Rozenberg and A. Salomaa. Handbook of Formal Languages: Volume 3 Beyond Words. Springer Science \& Business Media, 2012.
19. I. Schiering and W. Thomas. Counter-free automata, first-order logic, and starfree expressions extended by prefix oracles. Developments in Language Theory, II (Magdeburg, 1995), Worl Sci. Publishing, River Edge, NJ, pages 166-175, 1996.
20. B. Schneier. Attack trees. Dr. Dobb's journal, 24(12):21-29, 1999.
21. L. J. Stockmeyer. The complexity of decision problems in automata theory and logic. PhD thesis, Massachusetts Institute of Technology, 1974.
22. W. Thomas. Classifying regular events in symbolic logic. Journal of Computer and System Sciences, 25(3):360-376, 1982.
23. W. Thomas. An application of the ehrenfeucht-fraïssé game in formal language theory. Bull. Soc. Math. France, 16(1):1-21, 1984.
24. W. Wideł, M. Audinot, B. Fila, and S. Pinchinat. Beyond 2014: Formal methods for attack tree-based security modeling. ACM Computing Surveys (CSUR), 52(4):136, 2019.

## A Figure of $\tau_{e x_{2}}$ in Example 4



## B Proof of Theorem 1

We consider the lift binary relation between traces (Definition 5), a slight variant of the classic relation of super-word in language theory. We use the lift to define generators of a set of traces (see Definition 6) for $\tau$ an adt. We define (in Definition 7) $\operatorname{gen}(\tau)$, a finite set of traces, and we show in Proposition 3 that, if $\tau \in A D T_{1}$, then $\operatorname{gen}(\tau)$ is a set of generators of the semantics of $\tau$. We also show in Lemma 5 that an element of $\operatorname{gen}(\tau)$ has its size bounded by the number of leaves of $\tau$. Lemma 5 and Proposition 3 are enough to deduce the small model property of $A D T_{1}$, stated in Theorem 1 .

Definition 5. A trace $t$ is a lift of a trace $t^{\prime}=a_{1} \ldots a_{n}$, written $t^{\prime} \sqsubseteq t$, whenever $t \in \Sigma^{*} a_{1} \ldots \Sigma^{*} a_{n}$.

Notice that $\sqsubseteq$ is an order over $\Sigma^{*}$ and that lift relation is like the relation of super-word where the last letters of the respective traces are identical. It can be shown by induction of trees that every $\tau \in A D T_{0}$, the set $\operatorname{Traces}(\tau)$ is $\sqsubseteq$-upward closed, and is therefore infinite:

Lemma 4. Let $\tau \in A D T_{0}$. If $t \in \operatorname{Traces}(\tau)$, then for each trace $t^{\prime}$ such that $t \sqsubseteq t^{\prime}$, we have $t^{\prime} \in \operatorname{Traces}(\tau)$.

Proof. We show this result by induction over the shape of an adt.
If $\tau=\epsilon$, then there is no lift of $\varepsilon$. If $\tau$ is a non-empty leaf $\gamma$, the result holds since the lift relation preserves the last letter.

Now, we assume that the result holds for two adts $\tau_{1}$ and $\tau_{2}$ and we need to show the result for $\operatorname{OR}\left(\tau_{1}, \tau_{2}\right), \operatorname{SAND}\left(\tau_{1}, \tau_{2}\right)$ and $\operatorname{AND}\left(\tau_{1}, \tau_{2}\right)$.

The three cases are direct and similar, we only write the case $\tau=\operatorname{SAND}\left(\tau_{1}, \tau_{2}\right)$ : if $t \in \operatorname{Traces}(\tau)$ and $t \sqsubseteq t^{\prime}$ for a certain trace $t^{\prime}$, then we can write $t=t_{1} \cdot t_{2}$ with $t_{1} \in \operatorname{Traces}\left(\tau_{1}\right)$ and $t_{2} \in \operatorname{Traces}\left(\tau_{2}\right)$. From the structure of a lift, we can also write $t^{\prime}=t_{1}^{\prime} \cdot t_{2}^{\prime}$ such that $t_{1} \sqsubseteq t_{1}^{\prime}$ and $t_{2} \sqsubseteq t_{2}^{\prime}$, we can then conclude with our induction hypothesis.

From Lemma 4, it is immediate that any non-empty semantics of an adt in $A D T_{0}$ is necessarily infinite. We now turn to the notion of generators to establish the Small Model Property.

Definition 6. Let $L \subseteq \Sigma^{*}$. A set of generators of $L$ is a finite subset $G$ of $L$ with the following property: for each non-empty trace $t_{L} \in L$ there exists $t_{\text {gen }}=a_{1} \ldots a_{n} \in G$ such that we can write $t_{L}=w_{1} a_{1} \ldots w_{n} a_{n}$ and for each $t \in \Sigma^{*}$, if we can write $t=w_{1}^{\prime} a_{1} \ldots w_{1}^{\prime} a_{n}$ with $w_{1}^{\prime} a_{1} \sqsubseteq w_{1} a_{1}, \ldots, w_{n}^{\prime} a_{n} \sqsubseteq w_{n} a_{n}$, then $t \in L$.

Example 6. For the singleton set of propositions $\{p\}$, we consider the set of traces $L=p^{+} \backslash\{p p\}$. The set $\{p, p p p\}$ is a set of generators for $T$. However, the set $\{p\}$ is not because we have $p p p \in T$ and $p \sqsubseteq p p p$ but we also have $p \sqsubseteq p p \sqsubseteq p p p$ and $p p \notin T$.

Example 7. For the singleton set of propositions $\{p\}$, we let $a$ be the valuation that makes $p$ true and $b$ the valuation that makes $p$ false. Thus each trace is a word over the alphabet $\{a, b\}$. Now, remark that the set $(a b)^{+}$has no set of generators. Indeed, assume that $G \subseteq(a b)^{+}$is a set of generators. We write $(a b)^{N}$ for the largest element of $G$, thus $(a b)^{N} a b \in(a b)^{+} \backslash G$. However, for each $t_{\text {gen }} \in G$, we have $t_{\text {gen }} \sqsubseteq(a b)^{N} b \sqsubseteq(a b)^{N} a b$ and $(a b)^{N} b \notin(a b)^{+}$, which contradicts Definition 6 .

We now define the set $\operatorname{gen}(\tau)$ for each adt $\tau$ that will be our candidate to show that every of adts in $A D T_{1}$ has a set of generators.

Definition 7. Let $\tau \in A D T_{1}$. The set gen $(\tau)$ is inductively defined as follows:
$-\operatorname{gen}(\epsilon)=\emptyset$ and $\operatorname{gen}(\gamma)=\left\{a \in 2^{\text {Prop }}: a \models \gamma\right\}$,
$-\operatorname{gen}\left(\mathrm{OR}\left(\tau_{1}, \tau_{2}\right)\right)=\operatorname{gen}\left(\tau_{1}\right) \cup \operatorname{gen}\left(\tau_{2}\right)$,
$-\operatorname{gen}\left(\operatorname{SAND}\left(\tau_{1}, \tau_{2}\right)\right)=\operatorname{gen}\left(\tau_{1}\right) \cdot \operatorname{gen}\left(\tau_{2}\right)$,
$-\operatorname{gen}\left(\operatorname{CO}\left(\tau_{1}, \tau_{2}\right)\right)=\operatorname{gen}\left(\tau_{1}\right) \backslash \operatorname{Traces}\left(\tau_{2}\right)$,
$-\operatorname{gen}\left(\operatorname{AND}\left(\tau_{1}, \tau_{2}\right)\right)=\left(\operatorname{gen}\left(\tau_{1}\right) \bowtie \operatorname{gen}\left(\tau_{2}\right)\right) \cap \operatorname{Traces}\left(\operatorname{AND}\left(\tau_{1}, \tau_{2}\right)\right)$, where $\bowtie$ is the classic shuffle operator inductively defined as: $\varepsilon \bowtie t=t \bowtie \varepsilon=\{t\}$, $a_{1} t_{1} \bowtie a_{2} t_{2}=a_{1} \cdot\left(t_{1} \bowtie a_{2} t_{2}\right) \cup a_{2} \cdot\left(a_{1} t_{1} \bowtie t_{2}\right)$ and $a t_{1} \bowtie a t_{2}=a \cdot\left(t_{1} \bowtie\right.$ $\left.a t_{2}\right) \cup a \cdot\left(a t_{1} \bowtie t_{2}\right) \cup a \cdot\left(t_{1} \bowtie t_{2}\right)$.

We can prove the following Lemma 5 by induction over adts.
Lemma 5. For each trace $t \in \operatorname{gen}(\tau)$, its size $|t|$ is bounded by the number of leaves of $\tau$.

Proof. We show this result by induction over the shape of an adt.
If the adt is the empty leaf $\epsilon$, then there is no element in $\operatorname{gen}(\tau)$. If adt is a non-empty leaf $\gamma$ then all elements of $\operatorname{gen}(\tau)$ are of size 1 .

Now, we assume that the result holds for two adts $\tau_{1}$ and $\tau_{2}$ and we show the result for $\operatorname{gen}\left(\operatorname{OR}\left(\tau_{1}, \tau_{2}\right)\right)$, $\operatorname{gen}\left(\operatorname{SAND}\left(\tau_{1}, \tau_{2}\right)\right)$, $\operatorname{gen}\left(\operatorname{AND}\left(\tau_{1}, \tau_{2}\right)\right)$ and $\operatorname{gen}\left(\operatorname{CO}\left(\tau_{1}, \tau_{2}\right)\right)$. For the rest of the proof, we write $n_{1}$ the number of leaves of $\tau_{1}$ and $n_{2}$ the number of leaves of $\tau_{2}$.

By induction, if $t \in \operatorname{gen}\left(\operatorname{OR}\left(\tau_{1}, \tau_{2}\right)\right)$, then $|t| \leq M A X\left\{n_{1}, n_{2}\right\}$. If $t \in \operatorname{gen}\left(\operatorname{SaND}\left(\tau_{1}, \tau_{2}\right)\right)$ or $t \in \operatorname{gen}\left(\operatorname{AND}\left(\tau_{1}, \tau_{2}\right)\right)$, then $|t| \leq n_{1}+n_{2}$. Finally, if $t \in \operatorname{gen}\left(\operatorname{CO}\left(\tau_{1}, \tau_{2}\right)\right)$, then $|t| \leq n_{1}$. In each case, $|t| \leq n_{1}+n_{2}$ and $n_{1}+n_{2}$ is the number of leaves of the four studied adts, which concludes.

If $\tau$ has no nested countermeasures yields the following.
Proposition 3. When $\tau \in A D T_{1}$, the set $\operatorname{gen}(\tau)$ is a set of generators for Traces $(\tau)$.

Proof. We show this result by induction over the shape of an adt.
If $\tau=\epsilon$, then its semantics contains only the empty trace. Thus $\emptyset$ is a set of generators. If $\tau=\gamma$, then $\operatorname{Traces}(\gamma)=\left\{a_{1} \ldots a_{n} \in \Sigma^{*}: a_{n} \models \gamma\right\}=\Sigma^{*} \cdot \operatorname{gen}(\gamma)$, thus $\operatorname{gen}(\gamma) \subseteq \operatorname{Traces}(\gamma)$ and is a set of generators.

Now, for two adts $\tau_{1}$ and $\tau_{2}$, we assume that $\operatorname{gen}\left(\tau_{1}\right)$ and $\operatorname{gen}\left(\tau_{2}\right)$ are sets of generators and we show that $\operatorname{gen}\left(\operatorname{OR}\left(\tau_{1}, \tau_{2}\right)\right)$, $\operatorname{gen}\left(\operatorname{SAND}\left(\tau_{1}, \tau_{2}\right)\right)$ and $\operatorname{gen}\left(\operatorname{AND}\left(\tau_{1}, \tau_{2}\right)\right)$ are sets of generators for their corresponding adt.

First of all, for $\tau \in\left\{\operatorname{OR}\left(\tau_{1}, \tau_{2}\right), \operatorname{SAND}\left(\tau_{1}, \tau_{2}\right), \operatorname{AND}\left(\tau_{1}, \tau_{2}\right), \operatorname{CO}\left(\tau_{1}, \tau_{2}\right)\right\}$, we have $\operatorname{gen}(\tau) \subseteq \operatorname{Traces}(\tau)$. Indeed, for $\operatorname{gen}\left(\operatorname{OR}\left(\tau_{1}, \tau_{2}\right)\right)$ and $\operatorname{gen}\left(\operatorname{SAND}\left(\tau_{1}, \tau_{2}\right)\right)$, the composition rule is the same as the one used for $\operatorname{Traces}(\tau)$. Moreover, for $\operatorname{gen}\left(\operatorname{AND}\left(\tau_{1}, \tau_{2}\right)\right)$ and $\operatorname{gen}\left(\operatorname{CO}\left(\tau_{1}, \tau_{2}\right)\right)$, we intersect with $\operatorname{Traces}(\tau)$. Furthermore, from lemma5, each element of $\operatorname{gen}(\tau)$ is bounded, thus $\operatorname{gen}(\tau)$ is a finite set.

It remains to show that for each $t_{\tau} \in \operatorname{Traces}(\tau)$, we have that there exists $t_{\text {gen }}=a_{1} \ldots a_{n} \in \operatorname{gen}(\tau)$ such that, $t_{\tau}=t_{1} a_{1} \ldots t_{n} a_{n}$ and for each trace $t=t_{1}^{\prime} a_{1} \ldots t_{n}^{\prime} a_{n}$ with $t_{1}^{\prime} a_{1} \sqsubseteq t_{1} a_{1}, \ldots, t_{n}^{\prime} a_{n} \sqsubseteq t_{n} a_{n}$, we have $t \in \operatorname{Traces}(\tau)$. We distinguish the four adts operators:

$$
\begin{aligned}
& -\tau=\operatorname{OR}\left(\tau_{1}, \tau_{2}\right) \text {. If } t_{\tau} \in \operatorname{Traces}(\tau) \text {, then either } t_{\tau} \in \operatorname{Traces}\left(\tau_{1}\right) \text { or } t_{\tau} \in \\
& \operatorname{Traces}\left(\tau_{2}\right) \text {. With no loss of generality, we assume } t_{\tau} \in \operatorname{Traces}\left(\tau_{1}\right) \text {. Since } \\
& \operatorname{gen}\left(\tau_{1}\right) \text { is a set of generators for } \tau_{1} \text {, there exists } t_{\text {gen }}=a_{1} \ldots a_{n} \in \operatorname{gen}\left(\tau_{1}\right) \subseteq \\
& \operatorname{gen}(\tau) \text { such that } t_{\tau}=t_{1} a_{1} \ldots t_{n} a_{n} \text { and for each trace } t=t_{1}^{\prime} a_{1} \ldots t_{n}^{\prime} a_{n} \text { with } \\
& t_{1}^{\prime} a_{1} \sqsubseteq t_{1} a_{1}, \ldots, t_{n}^{\prime} a_{n} \sqsubseteq t_{n} a_{n} \text {, we have } t \in \operatorname{Traces}\left(\tau_{1}\right) \subseteq \operatorname{gen}(\tau) \text {. } \\
& -\tau=\overline{\operatorname{SAND}}\left(\tau_{1}, \tau_{2}\right) \text {. If } t_{\tau} \in \operatorname{Traces}(\tau) \text {, there exists } t_{\text {gen }}=a_{1} \ldots a_{n} \in \operatorname{gen}(\tau) \\
& \text { such that } t_{\tau}=t_{1} a_{1} \ldots t_{n} a_{n} \text {. Indeed, we can write } t_{\tau}=t_{1} \cdot t_{2} \text { with } t_{1} \in \\
& \operatorname{Traces}\left(\tau_{1}\right) \text { and } t_{2} \in \operatorname{Traces}\left(\tau_{2}\right) \text {, thus there exists } t_{\text {gen }_{1}}=a_{1} \ldots a_{i} \text { and } t_{\text {gen }_{2}}= \\
& a_{i+1} \ldots a_{n} \text { with } t_{\text {gen }_{1}} \in \operatorname{gen}\left(\tau_{1}\right) \text { and } t_{g e n_{2}} \in \operatorname{gen}\left(\tau_{2}\right) \text {. Therefore we can consider } \\
& t_{\text {gen }}=t_{\text {gen }_{1}} \cdot t_{\text {gen }}^{2} \text {. Moreover, the following holds: property for each trace } \\
& t=t_{1}^{\prime} a_{1} \ldots t_{n}^{\prime} a_{n} \text { with } t_{1}^{\prime} a_{1} \sqsubseteq t_{1} a_{1}, \ldots, t_{n}^{\prime} a_{n} \sqsubseteq t_{n} a_{n} \text {, we have } t \in \operatorname{Traces}\left(\tau_{1}\right) \subseteq \\
& \operatorname{gen}(\tau) \text {. Indeed we can simply write } t=t_{a} \cdot t_{b} \text { with } t_{a}=t_{1}^{\prime} a_{1} \ldots t_{i}^{\prime} a_{i} \text { and } t_{b}= \\
& t_{i+1}^{\prime} a_{i+1} \ldots t_{n} a_{n} \text {. So, } t_{a} \in \operatorname{Traces}\left(\tau_{1}\right) \text { and } t_{b} \in \operatorname{Traces}\left(\tau_{2}\right) \text { and we conclude } \\
& t \in \operatorname{Traces}(\tau) \text {. }
\end{aligned}
$$

$-\tau=\operatorname{AND}\left(\tau_{1}, \tau_{2}\right)$. First, let's remark that, for a trace $t_{\tau} \in \operatorname{Traces}(\tau)$, there exists $t_{1} \in \operatorname{Traces}\left(\tau_{1}\right)$ and $t_{2} \in \operatorname{Traces}\left(\tau_{2}\right)$ and $t_{\text {gen }}=a_{1} \ldots a_{n} \in\left(t_{1} \bowtie t_{2}\right)$ such that $t_{\tau}=t_{1} a_{1} \ldots t_{n} a_{n}$. Indeed, we can write $t_{\tau}=t_{a} \cdot t_{b}$ with either $t_{a} \in$ $\operatorname{Traces}\left(\tau_{1}\right)$ and $t_{\tau} \in \operatorname{Traces}\left(\tau_{2}\right)$ or $t_{a} \in \operatorname{Traces}\left(\tau_{2}\right)$ and $t_{\tau} \in \operatorname{Traces}\left(\tau_{1}\right)$. If we assume the former case (the latter is symmetrical), then there exists $t_{\text {gen }_{1}}=b_{1} \ldots b_{m} \in \operatorname{gen}\left(\tau_{1}\right)$ such that $t_{a}=u_{1} b_{1} \ldots u_{m} b_{m}$ and there exists $t_{g e n_{2}}=$ $c_{1} \ldots c_{p} \in \operatorname{gen}\left(\tau_{2}\right)$ such that $t_{\tau}=v_{1} c_{1} \ldots v_{p} c_{p}$. If we consider the trace $t_{g e n}$ as the trace obtained from $t_{\tau}$ by removing all valuations appearing neither in $t_{\text {gen }_{1}}$ nor in $t_{\text {gen }_{2}}$. Then, by construction: $t_{\text {gen }}=a_{1} \ldots a_{n} \in\left(t_{\text {gen }} \bowtie t_{\text {gen }_{2}}\right)$ and $t_{\tau}=t_{1} a_{1} \ldots t_{n} a_{n}$. The rest of the proof is similar to the SAND as we show that a trace $t=t_{1}^{\prime} a_{1} \ldots t_{n}^{\prime} a_{n}$ with $t_{1}^{\prime} a_{1} \sqsubseteq t_{1} a_{1}, \ldots, t_{n}^{\prime} a_{n} \sqsubseteq t_{n} a_{n}$ always can be written as $t=t_{1} \cdot t_{2}$ with $t_{1} \in \operatorname{Traces}\left(\tau_{1}\right)$ and $t \in \operatorname{Traces}\left(\tau_{2}\right)$.
$-\tau=\operatorname{CO}\left(\tau_{1}, \tau_{2}\right)$. For a trace $t_{\tau} \in \operatorname{Traces}(\tau)$, there exists $t_{\text {gen }}^{1} 10, ~ a_{1} \ldots a_{n} \in$ $\operatorname{gen}\left(\tau_{1}\right)$ such that, $t=t_{1}^{\prime} a_{1} \ldots t_{n}^{\prime} a_{n}$ with $t_{1}^{\prime} a_{1} \sqsubseteq t_{1} a_{1}, \ldots, t_{n}^{\prime} a_{n} \sqsubseteq t_{n} a_{n}$, we have $t \in \operatorname{Traces}\left(\tau_{1}\right)$. Moreover, if $t_{\tau} \in \operatorname{Traces}\left(\operatorname{Co}\left(\tau_{1}, \tau_{2}\right)\right)$, then $t_{\tau} \notin \operatorname{Traces}\left(\tau_{2}\right)$. Furthermore, if $\operatorname{CO}\left(\tau_{1}, \tau_{2}\right)$ is in $A D T_{1}$, then $\tau_{2} \in A D T_{0}$, thus from Lemma 4 , for each $t^{\prime}$ such that $t^{\prime} \sqsubseteq t_{\tau}$ we have $t^{\prime} \notin \operatorname{Traces}\left(\tau_{2}\right)$. This shows us that $t_{\text {gen }_{1}} \in \operatorname{Traces}\left(\operatorname{CO}\left(\tau_{1}, \tau_{2}\right)\right)$ and by taking $t_{\text {gen }}=t_{\text {gen }_{1}}$, the property we need to prove holds.

Remark 3. An adt $\tau \in A D T_{0}$ with a non-empty semantics always can be written as an OR over SAND over leaves. Indeed, for an adt $\tau \in A D T_{0}$ and its set of generator $\operatorname{gen}(\tau)=\left\{t_{1}, \ldots, t_{n}\right\}$, by combining Proposition 3 and Lemma 4 , we have: $\operatorname{Traces}(\tau)=\bigcup \Sigma^{*} a_{1} \ldots \Sigma^{*} a_{m}$. Therefore, for we can write $a_{1} \ldots a_{m} \in \operatorname{gen}(\tau)$ $\tau=\operatorname{OR}\left(\tau_{t_{1}}, \ldots, \tau_{t_{n}}\right)$, Where $\tau_{a_{1} \ldots a_{m}}=\operatorname{SAND}\left(\gamma_{1}, \ldots, \gamma_{m}\right)$ with $\gamma_{j}$ is the formula only satisfied by $a_{j}$.

By Proposition 3, for $\tau \in A D T_{1}$, if some non-empty trace $t_{\tau} \in \operatorname{Traces}(\tau)$, there is some $t_{\text {gen }} \in \operatorname{gen}(\tau)$ with $t_{g e n} \sqsubseteq t_{\tau}$ that is a witness. Together with Lemma 5, we conclude the proof of Theorem 1. Remark that, since $A D T_{0} \subseteq$ $A D T_{1}$, the result also holds for $A D T_{0}$.

Observe that our proof technique heavily relies on the notion of set of generators, which calls for the existence of a set of generators for adts above $A D T_{1}$. There is an adt in $A D T_{2}$ that shows it is hopeless: this adt specifies the language $(a b)^{+}$(see Section 5 but has no set of generators (see Example 7 ).

## C Complements of Section 4

Lemma 1. - For any $t \in \Sigma^{*}, t \models \psi_{\tau}$ iff $t \in \operatorname{Traces}(\tau)$.

- For any adt $\tau$, formula $\psi_{\tau}$ is of size exponential in $|\tau|$.

Proof. We will show this result by induction over the adt.
If $\tau=\epsilon$, then we have clearly that $\psi_{\epsilon}=\{\varepsilon\}$. Similarly, for a Boolean formula $\gamma$, we have that $\operatorname{Words}\left(\psi_{\gamma}\right)$ is all traces finishing with a valuation $a$ such that $a \models \gamma$. Thus, the property holds for leaves attack-defense trees.

Now, if we assume that the property is true for $\tau_{1}$ and $\tau_{2}$, then, since semantics for the or operator is a union of sets and semantics of $C$ operator is a difference of sets, the property for $\psi_{\mathrm{OR}\left(\tau_{1}, \tau_{2}\right)}$ and $\psi_{\mathrm{Co}\left(\tau_{1}, \tau_{2}\right)}$ trivially holds.

The formulas for AND and SAND operators are a little more difficult to understand since we need to distinguish whether an empty trace is in semantics of $\tau_{1}$ or $\tau_{2}$. Indeed, a $t$ is in semantics of $\operatorname{SAND}\left(\tau_{1}, \tau_{2}\right)$ if and only if we can find $t_{1}$ and $t_{2}$ such that $t_{1} \in \operatorname{Traces}\left(\tau_{1}\right)$ and $t_{2} \in \operatorname{Traces}\left(\tau_{2}\right)$ if $t_{1}$ is nonempty, then this is equivalent to say that there exists a position in the trace where we can cut as done in the first parenthesis of $\psi_{\operatorname{SAND}\left(\tau_{1}, \tau_{2}\right)}$. Otherwise, if $t_{1}$ is empty, then we handle this case with the use of $\psi_{\tau_{1}}^{\leq 0}$ in the second parenthesis of $\psi_{\operatorname{SAND}\left(\tau_{1}, \tau_{2}\right)}$. The idea for the AND operator is completely symmetrical except that we also need to consider the case where $\tau_{2}$ has the empty trace in its semantics.

Finally, it is clear that $\psi_{\tau}$ is of size exponential as subformulas for subtrees may be duplicated (for SAND and AND).

For the rest of this section, we often switch from an adt $\tau$ to its corresponding FO[ $[<]$-formulas $\psi_{\tau}$ as defined in Section 4 , always while assuming Lemma 1

Let us recall some basic results on $\Sigma_{k}$ and $\Pi_{k}$ hierarchies.
Remark 4 ([22], Lemma 2.4).
(a) The negation of a $\Sigma_{k}$-formula is equivalent to a $\Pi_{k}$-formula.
(b) A disjunction or conjunction of $\Sigma_{k}$-formulas is equivalent to a $\Sigma_{k}$ formula.
(c) A Boolean combination of $\Sigma_{k}$-formula is equivalent to a $\Sigma_{k+1}$-formula.
(d) The statements (a)-(c) hold in dual form for $\Pi_{k}$-formulas.

Lemma 2. 1. $A D T_{0} \subseteq \Sigma_{2} \cap \Pi_{2}$ - with an effective translation.
2. For every $k>0, A D T_{k} \subseteq \Sigma_{k+1}$ - with an effective translation.

Proof. Regarding Item 11 for both cases $A D T_{0} \subseteq \Sigma_{2}$ and $A D T_{0} \subseteq \Pi_{2}$, we construct for each $\tau \in A D T_{0}$ an $\mathrm{FO}[<]$-formula $\psi_{\tau}$ by induction over $\tau$ such that $\operatorname{Traces}(\tau)=\left\{t \in \Sigma^{*}: t=\psi_{\tau}\right\}$.

To prove that $A D T_{0} \subseteq \Sigma_{2}$, one can refer to the translation from ADT into $\mathrm{FO}[<]$, and easily verify that formula $\psi_{\tau} \in \Sigma_{2}$.

On the contrary, proving that $A D T_{0} \subseteq \Pi_{2}$ requires some work. First remark that $\psi_{\epsilon} \in \Pi_{1} \subseteq \Pi_{2}$. Moreover, recall that for a Boolean formula $\gamma$,

$$
\psi_{\gamma} \equiv \forall y \exists x(\neg(x<y) \wedge \hat{\gamma}(x))
$$

where $\widehat{\gamma}(x):=\bigvee_{a \models \gamma} a(x)$.
Indeed, both formulae hold for a trace $a_{1} \ldots a_{n}$ if, and only if, $a_{n} \models \gamma$.
Recall (Remark 3) that an arbitrary adt $\tau \in A D T_{0}$, can be equivalently written $\operatorname{OR}\left(\tau_{1}, \ldots, \tau_{n}\right)$ where $\tau_{1}, \ldots \tau_{n}$ are either the empty leaf $\epsilon$ or of the form $\operatorname{SAND}\left(\gamma_{1}, \ldots, \gamma_{m}\right)$. Consider the following $\Pi_{2}$-formula for $\operatorname{SAND}\left(\gamma_{1}, \ldots, \gamma_{m}\right)$ :

$$
\begin{aligned}
\theta_{\mathrm{SAND}\left(\gamma_{1}, \gamma_{2}, \ldots, \gamma_{m}\right)}:=\forall y \exists x_{1} \exists x_{2} \ldots \exists x_{m} \quad & x_{1}<x_{2}<\ldots<x_{m} \wedge y \leq x_{m} \wedge \\
& \hat{\gamma_{1}}\left(x_{1}\right) \wedge \hat{\gamma_{2}}\left(x_{2}\right) \wedge \ldots \wedge \hat{\gamma_{m}}\left(x_{m}\right)
\end{aligned}
$$

Clearly, for a word $w \in \Sigma^{*}$, we have $w \in \operatorname{Traces}\left(\operatorname{SAND}\left(\gamma_{1}, \ldots, \gamma_{m}\right)\right)$ if, and only if, $w \in \Sigma^{*} a_{1} \cdot \ldots \cdot \Sigma^{*} a_{m}$ with $a_{1} \models \gamma_{1}, \ldots, a_{m} \vDash \gamma_{m}$, which is exactly described by $\theta_{\operatorname{SAND}\left(\gamma_{1}, \gamma_{2}, \ldots, \gamma_{m}\right)} \in \Pi_{2}$. We now get back to the construction of the $\Pi_{2}$-formula for $\tau$, that is a disjunction of either $\psi_{\epsilon}$, or formulae of the form $\left.\theta_{\text {SAND }\left(\gamma_{1}, \gamma_{2}, \ldots, \gamma_{m}\right)}\right)$. Since they all belong to $\Pi_{2}$, so does their disjunction.

This concludes the proof of Item 1 .
For Item 2 of Lemma 2 we start with some notations: let $k>0$, for each $d \in \mathbb{N}$, we define $A D T_{k}(d) \subseteq A D T_{k}$ as follows.
$\left\{\begin{array}{l}A D T_{k}(0):=A D T_{k-1} \cup\left\{\operatorname{CO}\left(\tau_{1}, \tau_{2}\right): \tau_{1}, \tau_{2} \in A D T_{k-1}\right\} \\ A D T_{k}(d+1):=A D T_{k}(d) \cup\left\{\operatorname{OP}\left(\tau_{1}, \tau_{2}\right): \text { OP arbitrary, and } \tau_{1}, \tau_{2} \in A D T_{k}(d)\right\}\end{array}\right.$
We clearly have $\bigcup_{d \in \mathbb{N}} A D T_{k}(d)=A D T_{k}$.
We establish by a double induction over $k>0$ and $d \geq 0$ that for each $d \in \mathbb{N}$, $A D T_{k}(d) \subseteq \Sigma_{k+1}$, which clearly entails $A D T_{k} \subseteq \Sigma_{k+1}$.
$k=1$ : that is to show $A D T_{1}(d) \subseteq \Sigma_{2}$, for every $d \in \mathbb{N}$.
$d=0$ : Let adt $\tau \in A D T_{1}(0)$. If $\tau \in A D T_{0}$, it is immediate by Item 1 that $\operatorname{Traces}(\tau)$ is definable by a $\Sigma_{2}$ formula.
Otherwise $\tau=\operatorname{CO}\left(\tau_{1}, \tau_{2}\right)$ where $\tau_{1}$ and $\tau_{2} \in A D T_{0}$. Again, by Item 1 there exists formulas $\psi_{1} \in \Sigma_{2}$ and $\psi_{2} \in \Pi_{2}$ that characterize $\operatorname{Traces}\left(\tau_{1}\right)$ and $\operatorname{Traces}\left(\tau_{2}\right)$ respectively, so that, by the very definition of operator CO, formula $\psi_{1} \wedge \neg \psi_{2}$ characterizes $\operatorname{Traces}(\tau)$ and belongs to $\Sigma_{2}$ (see Remark 4).
$d>0$ : Let $\tau \in A D T_{1}(d)$; note that $\delta(\tau) \leq 1$. If $\tau \in A D T_{1}(d-1)$, we use the induction hypothesis over $d-1$ and we are done. Otherwise, $\tau=\mathrm{OP}\left(\tau_{1}, \tau_{2}\right)$.
Suppose OP $\in\{\mathrm{OR}, \mathrm{SAND}, \mathrm{AND}\}$, and because $\tau_{1}, \tau_{2} \in A D T_{1}(d-1)$, we know by induction over $d-1$ that there exist $\psi_{1}, \psi_{2} \in \Sigma_{2}$ that characterize $\operatorname{Traces}\left(\tau_{1}\right)$ and $\operatorname{Traces}\left(\tau_{2}\right)$ respectively. For $\mathrm{OP}=\mathrm{OR}$ (resp. $=$ SAND, AND), formula $\psi=\psi_{1} \vee \psi_{2}$ (resp. $=\exists x\left[\psi_{1}^{\leq x} \wedge \psi_{2}{ }^{>x}\right] \vee\left(\psi_{1}^{\leq 0} \wedge \psi_{2}\right)$, $\left.=\exists x\left[\left(\psi_{1}^{\leq x} \wedge \psi_{2}\right) \vee\left(\psi_{2} \leq x \wedge \psi_{1}\right)\right] \vee\left(\psi_{1}^{\leq 0} \wedge \psi_{2}\right) \vee\left(\psi_{2}{ }^{\leq 0} \wedge \psi_{1}\right)\right)$ characterises $\operatorname{Traces}\left(\mathrm{OP}\left(\tau_{1}, \tau_{2}\right)\right)$ and belongs to $\Sigma_{2}$ (see Remark 2 ).
Now, suppose $\mathrm{OP}=\mathrm{CO}$. Therefore, $\tau_{2} \in A D T_{0}(d-1)$ otherwise $\delta(\tau)=2$. By Item 11 we can assume that $\psi_{2} \in \Pi_{2}$ Now formula $\psi=\psi_{1} \wedge \neg \psi_{2}$ characterises $\operatorname{Traces}(\tau)$ and is indeed in $\Sigma_{2}$ (see Remark 4).
$k>1$ : that is to show $A D T_{k}(d) \subseteq \Sigma_{k+1}$, for every $d \in \mathbb{N}$.
$d=0$ : Let $\tau \in A D T_{k}(0)$. If $\tau \in A D T_{k-1}$, we resort to the induction hypothesis over $k$ to conclude since $\Sigma_{k} \subseteq \Sigma_{k+1}$. Otherwise, $\tau=\mathrm{CO}\left(\tau_{1}, \tau_{2}\right)$ with $\tau_{1}$ and $\tau_{2} \in A D T_{k-1}$. By induction over $k-1, \operatorname{Traces}\left(\tau_{1}\right)$ and $\operatorname{Traces}\left(\tau_{2}\right)$ can be equivalently characterized by $\Sigma_{k}$-formulas, say $\psi_{1}$ and $\psi_{2}$ respectively. Now the formula $\psi:=\psi_{1} \wedge \neg \psi_{2}$ characterizes $\operatorname{Traces}(\tau)$ and, by Remark 4 is clearly in $\Sigma_{k+1}$.
$d>0$ : Let $\tau \in A D T_{k}(d)$. If $\tau \in A D T_{k}(d-1)$, we use the induction hypothesis on $d-1$. Otherwise, $\tau=\mathrm{OP}\left(\tau_{1}, \tau_{2}\right)$ with $\tau_{1}, \tau_{2} \in A D T_{k}(d-1)$.

We can proceed in a way similar to what we did for the previous case $k=1, d>0$, by noticing that for the case where $\mathrm{OP}=\mathrm{CO}$, namely $\tau=\operatorname{Co}\left(\tau_{1}, \tau_{2}\right)$, the tree $\tau_{2} \in A D T_{k-1}$ otherwise $\tau$ would not belong to $A D T_{k}$.

Lemma 3. 1. $B_{\ell} \subseteq A D T_{2 \ell+2}$, and therefore $\Sigma_{\ell} \subseteq A D T_{2 \ell+2}$
2. $\Sigma_{1} \subseteq A D T_{0}$ - with an effective translation.

Proof. For Item 1, the proof is conducted by induction over ell. We start by showing that $B_{0} \subseteq A D T_{2}$. We recall that a language $L$ is in $B_{0}$ if it is finite or co-finite. We distinguish the two cases.

If $L$ is finite, we have that $L=\left\{t_{1}, \ldots, t_{n}\right\}=\left\{t_{1}\right\} \cup \ldots \cup\left\{t_{n}\right\}$ where $t_{1}, \ldots, t_{n} \in$ $\Sigma^{*}$. For a trace $t=a_{1}, \ldots, a_{n}$, we use the notation $\tau_{t}$ to express the adt $\operatorname{SAND}\left(\underline{a_{1}}, \ldots, \underline{a_{n}}\right)$. Clearly, $\operatorname{Traces}\left(\tau_{t}\right)=\{t\}$ and $\tau_{t} \in A D T_{1}$. Then if we consider the adt $\overline{\tau_{L}}=$ $\operatorname{OR}\left(\tau_{t_{1}}, \ldots, \tau_{t_{n}}\right)$, we have by construction that $\operatorname{Traces}\left(\tau_{L}\right)=L$. Moreover $\tau_{L} \in$ $A D T_{1}$ since an adt $\tau_{t}$ only uses non-nested co.

If $L$ is co-finite, then it can be written as $\Sigma \backslash L^{\prime}$ with $L^{\prime}$ a finite language. Thus, by definition of CO , if we consider $\tau_{L}=\operatorname{CO}\left(\top, \tau_{L^{\prime}}\right)$, we have $\operatorname{Traces}\left(\tau_{L}\right)=L$. We also have $\tau_{L} \in A D T_{2}$. Since $\tau_{L^{\prime}} \in A D T_{1}$. We concludes that $B_{0} \subseteq A D T_{2}$.

Now, for $k>0$, we assume $B_{\ell-1} \subseteq A D T_{2 \ell}$ and we show that $B_{\ell} \subseteq A D T_{2 \ell+2}$.
Let $L \in B_{\ell}$, so $L$ is a Boolean combination of languages of the form $L_{1} \cdot \ldots \cdot L_{n}$ with $L_{1}, \ldots, L_{n} \in B_{\ell-1}$. With no loss of generality, we can assume the Boolean combination written in disjunctive normal form where each conjunct is of the form $M_{1} \cap \ldots \cap M_{m}$, where the sets $M_{i}$ can be written either as $L_{1} \cdot \ldots \cdot L_{n}$ or as $\left(L_{1} \cdot \ldots \cdot L_{n}\right)^{c}$ with $L_{1}, \ldots, L_{n} \in B_{\ell-1}$. Moreover, We can rewrite each conjunct by using the equality $M_{1} \cap \ldots \cap M_{m}=\left(M_{1}^{c} \cup \ldots \cup M_{m}^{c}\right)^{c}$ to suppress all intersections of the Boolean expression, it results an expression only using union operators and complement operators. Furthermore, the maximal tower of complementary operator is 2. By induction hypothesis, we know that an expression $L_{1} \cdot \ldots \cdot L_{n}$ with $L_{1}, \ldots, L_{n} \in B_{\ell-1}$ is definable in $A D T_{2 \ell}$, indeed we consider the SAND of the adts associated with the languages $L_{1}, \ldots L_{n}$. Then, we can use the Boolean expression to construct the adt for $L$, where the union is replaced by the or and the complementation is replaced by NOT(.). Since the maximal tower of complementary operator of the Boolean expression is 2 , we need at most two more nested co to express the full Boolean expression. Therefore, the final adt is in $A D T_{2 \ell+2}$, which concludes.

Item 1 of Lemma 3 is an immediate corollary of the fact that $\Sigma_{k} \subseteq B_{\ell}$.
Regarding Item 2 of Lemma 3, let $\psi \in \Sigma_{1}$ a first-order formula of the form $\exists x_{1}, \ldots, \exists x_{n} \psi^{\prime}\left(x_{1}, \ldots, x_{n}\right)$ with $\psi^{\prime}$ quantifier-free. With no loss of generality, we consider $\psi^{\prime}=C_{1} \vee \ldots \vee C_{m}$ in disjunctive normal form, in other words $C_{1}, \ldots, C_{m}$ only use conjunctions of literals of the form $x_{i}<x_{j}$, or $a(x)$, or their negations. As expected, the adt corresponding to $\psi$ is of the form $\operatorname{OR}\left(\tau_{C_{1}}, \ldots, \tau_{C_{m}}\right)$, and we now explain how to build $\tau_{C}$.

If clause $C$ is not satisfiable, we associate adt $\perp$, otherwise we proceed as follows.

First, we show that with no loss of generality clause $C$ of $\psi^{\prime}$ specifies a linear order over $x_{1}, \ldots, x_{n}$ together with literals based on predicates $a\left(x_{i}\right)$. To do so, we
decompose the clause $C$, according to $C^{<} \wedge C^{a}$, where the first conjunct gathers all <-based literals and the second gathers the rest. Now, the sub-clause $C^{<}$ naturally induces a partial order between all free variables $x_{1}, \ldots, x_{n}$ of $\psi^{\prime}$. It is easy to see that clause $C^{<}$is equivalent to a disjunctive formula $\psi_{C}<$ where each disjunct describes a possible linearization of this partial order. For example, if $C^{<}=x<y \wedge \neg y<z$ then $\psi_{C}<=(x<y \wedge y=z) \vee(x<z \wedge z<y) \vee(z=$ $x \wedge x<y) \vee(z<x \wedge x<y)$, so that $C$ is equivalent to $\left(x<y \wedge y=z \wedge C^{a}\right) \vee(x<$ $\left.z \wedge z<y \wedge C^{a}\right) \vee\left(z=x \wedge x<y \wedge C^{a}\right) \vee\left(z<x \wedge x<y \wedge C^{a}\right)$, which concludes.

In clause $C$ (which now specifies a linear order of the variables), if $x_{i}=x_{j}$ with $i \neq j$, we replace all occurrences of $x_{j}$ by $x_{i}$ in $C$ and obtain a clause equivalent to $C$ of the form $x_{i_{1}}<\ldots<x_{i_{\ell}} \wedge P_{1}\left(x_{i_{1}}\right) \wedge \ldots \wedge P_{\ell}\left(x_{i_{\ell}}\right)$ where each $P_{k}\left(x_{i_{k}}\right)$ is the conjunction of literals of the form $a\left(x_{i_{k}}\right)$ and $\neg a\left(x_{i_{k}}\right)$. For convenience, we still write $C$ this clause.

We associate with clause $C$ the adt $\tau_{C}=\operatorname{SAND}\left(\gamma_{1}, \ldots, \gamma_{\ell}\right)\left\lceil\right.$ where $\gamma_{i}$ is made of the conjunction of the propositional formulas stemming from the valuations, or negation of valuations, that occur in $P_{k}\left(x_{i_{k}}\right)$. It can be shown that a trace $t \in \operatorname{Traces}\left(\tau_{C}\right)$ if, and only, if $t \equiv \exists x_{i_{1}} \ldots \exists x_{i_{\ell}} x_{i_{1}}<\ldots<x_{i_{\ell}} \wedge P_{1}\left(x_{i_{1}}\right) \wedge \ldots \wedge P_{\ell}\left(x_{i_{\ell}}\right)$.

## D Complements of Section 5

We start this section with some results and definitions to help us to prove Proposition 2. We then conduct the proof of Proposition2, and we finish with the proof of Lemma 8, a useful result to prove Proposition 1 .

Lemma 6. Let $t_{1}, t_{2}, t \in\{a, b\}^{*}$. The following assertions hold:

1. $\left\|t_{1} \cdot t_{2}\right\|=\left\|t_{1}\right\|+\left\|t_{2}\right\|$
2. If $\|t\|=k \geq 0$, then, for each $i \in\{0,1, \ldots, k\}$ there exists $t^{\prime} \preceq t$ such that $\left\|t^{\prime}\right\|=i$.
3. If $\|t\|=k>0$, then there exists $t_{1}$, $t_{2}$ such that $t=t_{1} \cdot a \cdot t_{2}$ with $\left\|t_{2}\right\|=0$ and for each $t_{2}^{\prime} \preceq t_{2}$, we have $\left\|t_{2}^{\prime}\right\| \geq 0$

Item 1 is trivial from the definition of $\|$.$\| , Item 2$ stems from the classic "Intermediate Value Theorem" in mathematics - that is applicable since extending a trace $t$ with a letter only increments or decrements $\|t\|$ by 1 . Item 3 is obtained by an application of Item 2 ,

We now start the proof of Proposition 2 .
Proposition 2. For each $k>0, W_{k} \in A D T_{k+1} \backslash A D T_{k-2}$.
Proof. Recall $W_{k} \notin A D T_{k-2}$ (page 11). We here focus on proving $W_{k} \notin A D T_{k+1}$.
We introduce companion languages of $W_{k}$, that were introduced in [23] namely $W_{k}{ }^{+}$and $W_{k}{ }^{-}$:
$-W_{k}^{+}:=\left\{w \in \Sigma^{*}\right.$ s.t. $\|w\|=k$ and for each $\left.w^{\prime} \preceq w, 0 \leq\left\|w^{\prime}\right\| \leq k\right\}$,
$-W_{k}{ }^{-}:=\left\{w \in \Sigma^{*}\right.$ s.t. $\|w\|=-k$ and for each $\left.w^{\prime} \preceq w,-k \leq\left\|w^{\prime}\right\| \leq 0\right\}$,

In [23], alternative definitions of $W_{k}, W_{k}{ }^{+}$and $W_{k}{ }^{-}$are provided:
$-W_{0}=W_{0}{ }^{+}=W_{0}{ }^{-}:=\{\varepsilon\} ;$
$-W_{k+1}:=\left(W_{k}^{+} a \Sigma^{*} \cap \Sigma^{*} b W_{k}{ }^{-}\right) \backslash\left(\Sigma^{*} a W_{k}{ }^{+} a \Sigma^{*} \cup \Sigma^{*} b W_{k}{ }^{-} b \Sigma^{*}\right) ;$
$-W_{k+1}^{+}:=\left(W_{k}{ }^{+} a \Sigma^{*} \cap \Sigma^{*} a W_{k}{ }^{+}\right) \backslash\left(\Sigma^{*} a W_{k}{ }^{+} a \Sigma^{*} \cup \Sigma^{*} b W_{k}{ }^{-} b \Sigma^{*}\right) ;$
$-W_{k+1}{ }^{-}:=\left(W_{k}{ }^{-} b \Sigma^{*} \cap \Sigma^{*} b W_{k}{ }^{-}\right) \backslash\left(\Sigma^{*} a W_{k}{ }^{+} a \Sigma^{*} \cup \Sigma^{*} b W_{k}{ }^{-} b \Sigma^{*}\right)$.
In the following, we may use the most convenient characterisation of $W_{k}$, $W_{k}{ }^{+}$and $W_{k}{ }^{-}$. Our proof relies on a third characterisation. We build languages $U_{k}, U_{k}{ }^{+}$and $U_{k}^{-}$where we eliminate the $\cap$ operator occurring in the recursive definitions of $W_{k}, W_{k}^{+}$and $W_{k}{ }^{-}$. For $w \in\{a, b\}^{*}$ we let its swap be the word $\operatorname{swap}(w)$ obtained by swapping $a$ and $b$ in $w$. We lift this operator to languages like usual: $\operatorname{swap}(L)=\{\operatorname{swap}(w): w \in L\}$. Note that $W_{k}{ }^{-}=\operatorname{swap}\left(W_{k}{ }^{+}\right)$.

We set:

$$
\begin{aligned}
& -U_{0}=U_{0}^{+}=U_{0}^{-}:=\{\varepsilon\} ; \\
& -U_{k+1}:=\left(U_{k}^{+} a \Sigma^{*} b U_{k}^{-}\right) \backslash\left(\Sigma^{*} a U_{k}{ }^{+} a \Sigma^{*} \cup \Sigma^{*} b U_{k}^{-} b \Sigma^{*}\right) \\
& -U_{k+1}^{+}:=\left(\left(U_{k}^{+} a \Sigma^{*} a U_{k}^{+}\right) \cup\left(\bigcup_{i \leq k} U_{i} a U_{k}^{+}\right)\right) \backslash\left(\Sigma^{*} a U_{k}{ }^{+} a \Sigma^{*} \cup \Sigma^{*} b U_{k}^{-} b \Sigma^{*}\right) ; \\
& -U_{k+1}^{-}:=\operatorname{swap}\left(U_{k+1}{ }^{+}\right)
\end{aligned}
$$

Lemma 7. For every $k \geq 0$, we have $W_{k}=U_{k}, W_{k}{ }^{+}=U_{k}{ }^{+}$and $W_{k}{ }^{-}=U_{k}{ }^{-}$.
Proof. The proof is conducted by induction over $k$. By definition, $W_{0}=U_{0}$, $W_{0}{ }^{+}=U_{0}{ }^{+}$and $W_{0}{ }^{-}=U_{0}{ }^{-}$. For the rest of the proof, we fix $k>0$.

- We start to show $W_{k+1}=U_{k+1}$. Namely, by replacing $U_{k}$ (respectively $U_{k}{ }^{+}$, $U_{k}{ }^{-}$) by $W_{k}$ (respectively $W_{k}^{+}, W_{k}{ }^{-}$) that:

$$
\begin{aligned}
& \left(W_{k}^{+} a \Sigma^{*} \cap \Sigma^{*} b W_{k}^{-}\right) \backslash\left(\Sigma^{*} a W_{k}^{+} a \Sigma^{*} \cup \Sigma^{*} b W_{k}^{-} b \Sigma^{*}\right) \\
& \quad=\left(W_{k}^{+} a \Sigma^{*} b W_{k}^{-}\right) \backslash\left(\Sigma^{*} a W_{k}^{+} a \Sigma^{*} \cup \Sigma^{*} b W_{k}^{-} b \Sigma^{*}\right)
\end{aligned}
$$

It is enough to show that $\left(W_{k}^{+} a \Sigma^{*} \cap \Sigma^{*} b W_{k}{ }^{-}\right)=\left(W_{k}^{+} a \Sigma^{*} b W_{k}{ }^{-}\right)$for the equality to hold. Because $\left(W_{k}{ }^{+} a \Sigma^{*} \cap \Sigma^{*} b W_{k}{ }^{-}\right) \supseteq\left(W_{k}{ }^{+} a \Sigma^{*} b W_{k}{ }^{-}\right)$is clear, we focus on showing $\left(W_{k}{ }^{+} a \Sigma^{*} \cap \Sigma^{*} b W_{k}{ }^{-}\right) \subseteq\left(W_{k}{ }^{+} a \Sigma^{*} b W_{k}{ }^{-}\right)$. Let $w$ be in $W_{k}{ }^{+} a \Sigma^{*} \cap \Sigma^{*} b W_{k}^{-}$. As $w \in W_{k}^{+} a \Sigma^{*}$, we can write $w=u_{1} \cdot a \cdot u_{2}$ with $u_{1} \in W_{k}^{+}$and let $n$ be the position of the distinguished $a$ occurrence in $w$. As we also have $w \in \Sigma^{*} b W_{k}{ }^{-}$, we can write $w=v_{2} \cdot b \cdot v_{1}$ with $v_{1} \in W_{k}^{-}$ and let $m$ be the position of the distinguished $b$ occurrence in $w$. Clearly $n \neq m$. Moreover, we cannot have $m<n$. Indeed, if $m<n$ we can write $w=v_{2} \cdot b \cdot w^{\prime} \cdot a \cdot u_{2}$ with $v_{2} \cdot b \cdot w^{\prime}=u_{1} \in W_{k}^{+}$and $w^{\prime} \cdot a \cdot u_{2}=v_{1} \in W_{k}^{-}$. Since $v_{2}$ is a prefix of $u_{1} \in W_{k}^{+}$, we have $\left\|v_{2}\right\| \leq k$, and therefore $\left\|v_{2} \cdot b\right\| \leq k-1$ (by Item 1 of Lemma 6). Moreover, since $w^{\prime}$ is a prefix of $v_{1} \in W_{k}{ }^{-}$, we have $\left\|w^{\prime}\right\| \leq 0$. We conclude that $\left\|v_{2} \cdot b \cdot w^{\prime}\right\| \leq k-1$ (by Item 1 of Lemma 6), which contradicts $v_{2} \cdot b \cdot w^{\prime} \in W_{k}{ }^{+}$. Therefore $n<m$, and $w \in\left(W_{k}{ }^{+} a \Sigma^{*} b W_{k}{ }^{-}\right)$, which concludes.

- We now prove that $W_{k+1}{ }^{+}=U_{k+1}{ }^{+}$. Because the inclusion $W_{k+1}{ }^{+} \supseteq U_{k+1}{ }^{+}$ is clear, we focus on $W_{k+1}{ }^{+} \subseteq U_{k+1}{ }^{+}$. Namely, by replacing $U_{k}$ (respectively $U_{k}{ }^{+}, U_{k}{ }^{-}$) by $W_{k}$ (respectively $W_{k}{ }^{+}, W_{k}^{-}$) that:

$$
\begin{aligned}
\left(W_{k}{ }^{+} a \Sigma^{*} \cap \Sigma^{*} a W_{k}^{+}\right) \backslash\left(\Sigma^{*} a W_{k}{ }^{+} a \Sigma^{*} \cup \Sigma^{*} b W_{k}{ }^{-} b \Sigma^{*}\right) \subseteq \\
\left.\left(W_{k}+a \Sigma^{*} a W_{k}^{+}\right) \cup\left(\bigcup_{i \leq k} W_{i} a W_{k}{ }^{+}\right)\right) \backslash\left(\Sigma^{*} a W_{k}{ }^{+} a \Sigma^{*} \cup \Sigma^{*} b W_{k}^{-} b \Sigma^{*}\right)
\end{aligned}
$$

Let $w \in W_{k+1}{ }^{+}$. As, on the one hand, $w \in W_{k}{ }^{+} a \Sigma^{*}$, we can write $w=$ $u_{1} \cdot a \cdot u_{2}$ with $u_{1} \in W_{k}^{+}$and let $n$ be the position of the distinguished $a$ occurrence in $w$. As, on the other hand, we also have $w \in \Sigma^{*} a W_{k}{ }^{+}$, we can write $w=v_{2} \cdot a \cdot v_{1}$ with $v_{1} \in W_{k}{ }^{+}$and let $m$ be the position of the distinguished $a$ occurrence in $w$.
First of all, $n \neq m$ : indeed, if $n=m$, we can write $w=u_{1} \cdot a \cdot v_{1}$ with $u_{1}, v_{1} \in W_{k}{ }^{+}$. Since $\left\|u_{1} \cdot a\right\|=k+1$, we can write $u_{1} \cdot a=q_{1} \cdot a \cdot q_{2}$ with $q_{2} \in W_{k}{ }^{+}$. Since $v_{1} \in W_{k}{ }^{+}$, we can write $v_{1}=a \cdot v_{1}^{\prime}$. So that $w=$ $q_{1} \cdot a \cdot q_{2} \cdot a \cdot v_{1}^{\prime} \in\left(\Sigma^{*} a W_{k}^{+} a \Sigma^{*}\right)$ which entails $w \notin W_{k+1}$, leading to a contradiction.
We distinguish the two remaining cases.

- If $n<m$, then we have $w \in\left(W_{k}{ }^{+} a \Sigma^{*} a W_{k}{ }^{+}\right)$, which entails $w \in U_{k+1}{ }^{+}$.
- If $m<n$, we can write $u_{1}=p_{1} \cdot a \cdot p_{2} \in W_{k}{ }^{+}$and $v_{1}=p_{2} \cdot a \cdot p_{3} \in W_{k}{ }^{+}$, where $w=p_{1} \cdot a \cdot p_{2} \cdot a \cdot p_{3}$. Notice that $\left\|p_{1}\right\| \geq 0$, as $u_{1} \in W_{k}{ }^{+}$, but we show that $\left\|p_{1}\right\|=0$ : if $\left\|p_{1}\right\|>0$, we can rewrite $p_{1}=q_{1} \cdot a \cdot q_{2}$ with $\left\|q_{2}\right\|=0$ and for each $q_{2}^{\prime}$ prefix of $q_{2}$, we have $\left\|q_{2}^{\prime}\right\| \geq 0$ (by Item 3 of Lemma (6). Moreover, $\left\|a \cdot p_{2} \cdot a \cdot p_{3}\right\|=k+1$ (by Item 1 of Lemma 6 since $\left\|p_{2} \cdot a \cdot p_{3}\right\|=k$, thus we can write $a \cdot p_{2} \cdot a \cdot p_{3}=r_{1} \cdot a \cdot r_{2}$ with $r_{1} \in W_{k}{ }^{+}$(by Item 2 of Lemma 66). Therefore, $w=q_{1} \cdot a \cdot q_{2} \cdot r_{1} \cdot a \cdot r_{2}$ with $q_{2} \cdot r_{1} \in W_{k}^{+}$(by Item 1 of Lemma 6 6 , so that $w \in\left(\Sigma^{*} a W_{k}{ }^{+} a \Sigma^{*}\right)$, which contradicts $w \in W_{k+1}$.
Since $\left\|p_{1}\right\|=0$, we have $p_{1} \in \bigcup_{i \leq k} W_{i}$, and we conclude $w=p_{1} \cdot a$. $v_{1} \in \bigcup_{i \leq k} W_{i} a W_{k}^{+}$. Also because $w \in W_{k+1}{ }^{+}, w \notin\left(\Sigma^{*} a W_{k}{ }^{+} a \Sigma^{*} \cup\right.$ $\left.\Sigma^{*} b W_{k}{ }^{-} b \Sigma^{*}\right)$ and we obtain $w \in U_{k+1}{ }^{+}$, which concludes.
- Finally, $W_{k+1}{ }^{-}=U_{k+1}{ }^{-}$because $W_{k+1}{ }^{-}=\operatorname{swap}\left(W_{k+1}{ }^{+}\right)$and $U_{k+1}{ }^{-}=$ $\operatorname{swap}\left(U_{k+1}{ }^{+}\right)$and we have already shown $W_{k+1}{ }^{+}=U_{k+1}{ }^{+}$.

This conclude the proof of Lemma 7
We use Lemma 7 to construct three adts $\mu_{k}, \mu_{k}^{+}$and $\mu_{k}^{-}$such that $\operatorname{Traces}\left(\mu_{k}\right)=$ $W_{k}, \operatorname{Traces}\left(\mu_{k}^{+}\right)=W_{k}^{+}$and $\operatorname{Traces}\left(\mu_{k}^{-}\right)=W_{k}^{-}, \delta\left(\mu_{k}\right)=\delta\left(\mu_{k}^{+}\right)=\delta\left(\mu_{k}^{-}\right)=$ $k+1$ which achieves the proof of Proposition 2. The proof is conducted by induction over $k \geq 1$.

To capture $W_{1}$, we propose the following adt depicted in fig. 2a;

$$
\mu_{1}:=\operatorname{CO}(b, \operatorname{OR}(\underline{b}\lceil,\rceil \operatorname{CO}(|\geq 2|, \operatorname{AND}(a\lceil, b\lceil ))\lceil )),
$$

We prove $\operatorname{Traces}\left(\mu_{1}\right)=(a b)^{+}=W_{1}$ : we point to Example 2 for the semantics for $\operatorname{ALL}, \operatorname{NOT}(\tau)$ and $\rceil \tau\lceil$. First of all, $\operatorname{And}(a\lceil, b\lceil )$ defines the set of all traces with at least one occurrence of $a$ and one occurrence of $b$. Thus $\operatorname{Traces}\left(\mathrm{CO}\left(|\geq 2|, \operatorname{AND}(a\lceil, b\lceil )))=a a^{+} \cup b b^{+}\right.\right.$. Write $\tau:=\mathrm{CO}(|\geq 2|, \operatorname{AND}(a\lceil, b\lceil ))$. So, Traces $( \rceil \tau\left)=\Sigma^{*} .\left(a a^{+} \cup b b^{+}\right) . \Sigma^{*}=\Sigma^{*} . a a \cdot \Sigma^{*} \cup \Sigma^{*} . b b . \Sigma^{*}\right.$. On this basis, $\operatorname{Traces}\left(\mu_{1}\right)=\operatorname{Traces}\left(\mathrm{CO}(b, \operatorname{OR}(\underline{b}\lceil,\rceil \tau\lceil )))=\Sigma^{*} b \backslash\left(b \Sigma^{*} \cup \Sigma^{*} . a a \cdot \Sigma^{*} \cup \Sigma^{*} \cdot b b \cdot \Sigma^{*}\right)=\right.$ $(a b)^{+}=W_{1}$.

We now compute $\delta\left(\mu_{1}\right)$ since $\delta(\tau)=1$, we have $\delta\left(\mu_{1}\right)=\delta(\operatorname{CO}(b, \operatorname{OR}(\underline{b}\lceil,\rceil \tau\lceil )))=$ $\max \{0, \max \{1,1\}+1\}=2$.

Similarly, we define:

$$
\begin{aligned}
& -\mu_{1}^{+}:=\operatorname{co}(a, \operatorname{OR}(\underline{b}\lceil,\rceil \operatorname{co}(|\geq 2|, \operatorname{AND}(a\lceil, b\lceil )) \Gamma)) ; \\
& -\mu_{1}^{-}:=\operatorname{co}(b, \operatorname{OR}(\underline{a}\lceil,\rceil \operatorname{co}(|\geq 2|, \operatorname{AND}(a\lceil, b\lceil )) \Gamma)) .
\end{aligned}
$$

As done for $\mu_{1}$, one can verify that $\operatorname{Traces}\left(\mu_{1}^{+}\right)=(a b)^{*} a=W_{1}^{+}$and $\operatorname{Traces}\left(\mu_{1}^{-}\right)=$ $b(a b)^{*}=W_{1}^{-}$and that $\delta\left(\mu_{1}^{+}\right)=\delta\left(\mu_{1}^{-}\right)=2$.

We make use of $U_{k}, U_{k}^{+}$and $U_{k}^{-}$to inductively define $\mu_{k}, \mu_{k}^{+}$and $\mu_{k}^{-}$. For readability, we introduce for $k \geq 2$ the subtrees $\xi_{k}$ and $\zeta_{k}$ and we extend the definition of $\operatorname{swap}($.$) to adts by applying the swap to leaves.$

$$
\begin{gathered}
\xi_{k}:=\operatorname{OR}\left(\operatorname { S A N D } \left(a, \mu_{k}^{+}, \underline{a}\lceil ), \operatorname{SAND}\left(b, \mu_{k}^{-}, \underline{b}\lceil )\right)\right.\right. \\
\zeta_{k}:=\operatorname{OR}\left(\operatorname{SAND}\left(\mu_{1}, \underline{a}, \mu_{k}^{+}\right), \operatorname{sAND}\left(\mu_{2}, \underline{a}, \mu_{k}^{+}\right), \ldots, \operatorname{SAND}\left(\mu_{k}, \underline{a}, \mu_{k}^{+}\right)\right)
\end{gathered}
$$

And $\mu_{k}, \mu_{k}^{+}$and $\mu_{k}^{-}$are defined by:

```
\(-\mu_{k}:=\operatorname{CO}\left(\operatorname{SAND}\left(\mu_{k-1}^{+}, \underline{a}, \operatorname{ALL}, \underline{b}, \mu_{k-1}^{-}\right), \xi_{k-1}\right)\)
\(-\mu_{k}^{+}:=\operatorname{CO}\left(\operatorname{OR}\left(\operatorname{SAND}\left(\mu_{k-1}^{+}, \underline{a}, \operatorname{ALL}, \underline{a}, \mu_{k-1}^{+}\right), \zeta_{k-1}\right), \xi_{k-1}\right)\)
\(-\mu_{k}^{-}:=\operatorname{swap}\left(\mu_{k}^{+}\right)\)
```

The adts $\mu_{k}, \mu_{k}^{+}$and $\mu_{k}^{-}$are built in such a way that a direct application of adt semantics yields $W_{k}=\operatorname{Traces}\left(\mu_{k}\right), W_{k}{ }^{+}=\operatorname{Traces}\left(\mu_{k}^{+}\right)$and $W_{k}{ }^{-}=$ $\operatorname{Traces}\left(\mu_{k}^{-}\right)$.

It remains to show that $\delta\left(\mu_{k}\right)=\delta\left(\mu_{k}^{+}\right)=\delta\left(\mu_{k}^{-}\right)=k+1$ assuming $\delta\left(\mu_{k-1}\right)=$ $\delta\left(\mu_{k-1}^{+}\right)=\delta\left(\mu_{k-1}^{-}\right)=k$. By definition, $\delta\left(\mu_{k}\right)=\max \left\{\delta\left(\tau_{1}\right), \delta\left(\xi_{k-1}\right)+1\right\}$ with $\tau_{1}=$ $\operatorname{SAND}\left(\mu_{k-1}^{+}, \underline{a}, \operatorname{ALL}, \underline{b}, \mu_{k-1}^{-}\right)$. By using Example 5, we have $\delta\left(\tau_{1}\right)=\max \left\{\delta\left(\mu_{k-1}^{+}\right)\right.$, $\left.1,0,1, \delta\left(\mu_{k-1}^{-}\right)\right\}$and $\delta\left(\xi_{k-1}\right)=\max \left\{0, \delta\left(\mu_{k-1}^{+}\right), 1,0, \delta\left(\mu_{k-1}^{-}\right), 1\right\}$. Applying the induction hypothesis over $\delta\left(\mu_{k-1}^{+}\right)=\delta\left(\mu_{k-1}^{-}\right)=k$, we obtain: $\delta\left(\tau_{1}\right)=\max \{k, 1,0,1, k\}=$ $k$ and $\delta\left(\tau_{2}\right)=\max \{0, k, 1,0, k, 1\}=k$, thus $\delta\left(\mu_{k}\right)=\max \{k, k+1\}=k+1$, which concludes. Similarly, we can establish $\delta\left(\mu_{k}^{+}\right)=\delta\left(\mu_{k}^{-}\right)=k+1$.

We have shown $W_{k}=\operatorname{Traces}\left(\mu_{k}\right)$ with $\delta\left(\mu_{k}\right)=k+1$ which concludes the proof of Proposition 2 .

Lemma 8. If $A D T_{k_{0}}=A D T_{k_{0}+1}$ for some $k_{0}>0$, then all $A D T_{k}$ collapse from $k_{0}$.

Proof. Let $k>0$ be such that for all $\tau \in A D T_{k+1}$, we have that $\tau$ is $A D T_{k^{-}}$ definable. Given $\tau_{k+2} \in A D T_{k+2}$, we know that for each subtree of the form $\operatorname{CO}\left(\tau_{1}, \tau_{2}\right)$ in $\tau_{k+2}$, we have that $\tau_{2} \in A D T_{k+1}$. Therefore $\tau_{2}$ is $A D T_{k}$-definable. Hence there exists $\tau_{2}^{\prime} \in A D T_{k}$ such that $\tau_{2} \equiv \tau_{2}^{\prime}$. By replacing $\tau_{2}$ by $\tau_{2}^{\prime}$ in $\tau_{k+2}$, we do not change its semantics. By applying this procedure over all co operator of minimal depth (ie. CO operator having no CO in their ancestors), we obtain $\tau_{k+1}^{\prime} \in A D T_{k+1}$ such that $\tau_{k+1} \equiv \tau_{k+2}^{\prime}$. By hypothesis, we know that $\tau_{k+1}^{\prime}$ is $A D T_{k}$-definable. This implies that $\tau_{k+1}$ is also $A D T_{k}$-definable. We can then extend this result for each $l>k$ by induction.

## E Complements of Section 6

Regarding the upper-bound complexity of $A D T$-memb (Table 11), we present here an alternating algorithm using the following logarithmic space (hence a Ptime complexity for $A D T$-memb) Algorithm 1 . Remark that our algorithm can be extended to allow arbitrary EREs (with the Kleene star) as inputs, but this is out of the scope of the paper.

```
Algorithm \(1 \operatorname{Memb}(\tau, t)\)
```

Algorithm $1 \operatorname{Memb}(\tau, t)$
Input: $\tau$ an adt and $t$ a trace
Input: $\tau$ an adt and $t$ a trace
Output: True if $t \in \operatorname{Traces}(\tau)$, False otherwise.
Output: True if $t \in \operatorname{Traces}(\tau)$, False otherwise.
switch $(\tau)$
switch $(\tau)$
case $\epsilon$ :
case $\epsilon$ :
return $(t=\varepsilon)$
return $(t=\varepsilon)$
case $\tau=\gamma$ :
case $\tau=\gamma$ :
return $(\operatorname{last}(t) \models \gamma)$
return $(\operatorname{last}(t) \models \gamma)$
case $\tau=O P\left(\tau^{\prime}\right)$ :
case $\tau=O P\left(\tau^{\prime}\right)$ :
return $\operatorname{Memb}\left(\tau^{\prime}, t\right)$
return $\operatorname{Memb}\left(\tau^{\prime}, t\right)$
case $\tau=\mathrm{OR}\left(\tau_{1}, \ldots, \tau_{n}\right)$ :
case $\tau=\mathrm{OR}\left(\tau_{1}, \ldots, \tau_{n}\right)$ :
( $\exists$ ) guess $\tau^{\prime} \in\left\{\tau_{1}, \ldots, \tau_{n}\right\}$
( $\exists$ ) guess $\tau^{\prime} \in\left\{\tau_{1}, \ldots, \tau_{n}\right\}$
return $\operatorname{Memb}\left(\tau^{\prime}, t\right)$
return $\operatorname{Memb}\left(\tau^{\prime}, t\right)$
case $\tau=\operatorname{SAND}\left(\tau_{1}, \ldots, \tau_{n}\right)$ :
case $\tau=\operatorname{SAND}\left(\tau_{1}, \ldots, \tau_{n}\right)$ :
( $\exists$ ) guess $i \in\{1, \ldots,|t|\}$
( $\exists$ ) guess $i \in\{1, \ldots,|t|\}$
$(\forall)$ guess test $\in\{$ first, others $\}$
$(\forall)$ guess test $\in\{$ first, others $\}$
if test $=$ first then
if test $=$ first then
return $\operatorname{Memb}\left(\tau_{1}, t_{1} \ldots t_{i}\right)$
return $\operatorname{Memb}\left(\tau_{1}, t_{1} \ldots t_{i}\right)$
else
else
return
return
$\operatorname{Memb}\left(\operatorname{SAND}\left(\tau_{2}, \ldots, \tau_{n}\right), t_{i+1 \ldots t_{|t|}}\right)$
$\operatorname{Memb}\left(\operatorname{SAND}\left(\tau_{2}, \ldots, \tau_{n}\right), t_{i+1 \ldots t_{|t|}}\right)$
end if
end if
20: case $\tau=\operatorname{AND}\left(\tau_{1}, \ldots, \tau_{n}\right)$ :
20: case $\tau=\operatorname{AND}\left(\tau_{1}, \ldots, \tau_{n}\right)$ :
21: ( $\exists$ ) guess (i,a) $\in\{1, \ldots,|t|\} \times$
21: ( $\exists$ ) guess (i,a) $\in\{1, \ldots,|t|\} \times$
$\{1, \ldots, n\}$ ( $)$ guess test $\in$
$\{1, \ldots, n\}$ ( $)$ guess test $\in$
\{first,others $\}$
\{first,others $\}$
if test $=$ first then
if test $=$ first then
return $\operatorname{Memb}\left(\tau_{a}, t_{1} \ldots t_{i}\right)$
return $\operatorname{Memb}\left(\tau_{a}, t_{1} \ldots t_{i}\right)$
else
else
return
return
$\operatorname{Memb}\left(\operatorname{AND}\left(\tau_{1}, \ldots, \tau_{i-1}, \tau_{i+1}, \ldots, \tau_{n}\right), t\right)$
$\operatorname{Memb}\left(\operatorname{AND}\left(\tau_{1}, \ldots, \tau_{i-1}, \tau_{i+1}, \ldots, \tau_{n}\right), t\right)$
end if
end if
case $\tau=\mathrm{CO}\left(\tau_{1}, \tau_{2}\right)$ :
case $\tau=\mathrm{CO}\left(\tau_{1}, \tau_{2}\right)$ :
$(\forall)$ guess test $\in\{$ first, second $\}$
$(\forall)$ guess test $\in\{$ first, second $\}$
if test $=$ first then
if test $=$ first then
return $\operatorname{Memb}\left(\tau_{1}, t\right)$
return $\operatorname{Memb}\left(\tau_{1}, t\right)$
else
else
return $\neg \operatorname{Memb}\left(\tau_{2}, t\right)$
return $\neg \operatorname{Memb}\left(\tau_{2}, t\right)$
end if
end if
end switch

```
    end switch
```

Proposition 4. Algorithm 1 solves ADT-MEMB in logarithmic space.
Proof. We start by showing that Algorithm 1 runs in logarithmic space, then we prove its correctness.

Since at each recursive call we only need to recall over which factor of $t$ we are computing (in constant space) and over which part of $\tau$ (in constant space) we are pursuing the computation, Algorithm 1 runs in logarithmic space.

Regarding the correctness of Algorithm 1, we conduct a proof by induction over $\tau$.

The cases where $\tau$ is a leaf (lines 1 to 7 ) are correct by definition. For the case where $\tau=O P\left(\tau^{\prime}\right)$ (line 9), then, Algorithm 1 is correct too since $\operatorname{Traces}(\tau)=$ Traces $\left(\tau^{\prime}\right)$.

If $\tau=\mathrm{OR}\left(\tau_{1}, \ldots, \tau_{n}\right)$, then the case lines 13 to 16 are correct too since by definition, $t \in \operatorname{Traces}(\tau)$ if and only if one can find $i \in\{1, \ldots, n\}$ such that $t \in \operatorname{Traces}\left(\right.$ att $\left._{i}\right)$.

If $\tau=\operatorname{SAND}\left(\tau_{1}, \ldots, \tau_{n}\right)$, then, by associativity, $\tau \equiv \operatorname{SAND}\left(\tau_{1}, \operatorname{SAND}\left(\tau_{2}, \ldots, \tau_{n}\right)\right)$. Moreover, we have that $t=a_{1} \ldots a_{m} \in \operatorname{Traces}(\tau)$ if and only if one can find $i \in\{1, \ldots, m\}$ such that $a_{1} \ldots a_{i} \in \tau_{1}$ and $a_{i+1} \ldots a_{m} \in \operatorname{SAND}\left(\tau_{2}, \ldots, \tau_{n}\right)$, which is what is done in lines 18-26.

If $\tau=\operatorname{AND}\left(\tau_{1}, \ldots, \tau_{n}\right)$, then, by commutativity and associativity for each $1 \leq j \leq n$ we have $\tau \equiv \operatorname{AND}\left(\tau_{j}, \operatorname{AND}\left(\tau_{1}, \ldots, \tau_{j-1}, \tau_{j+1}, \ldots, \tau_{n}\right)\right)$. Moreover, if a trace $t \in \operatorname{AND}\left(\tau_{1}, \ldots, \tau_{n}\right)$, we can always choose $j$ such that there is $1 \leq k \leq n$, with $k \neq j$ and $t \in \operatorname{Traces}\left(\tau_{k}\right)$. Therefore $t$ is in $\operatorname{Traces}(\tau)$ if and only if $t$ has a prefix in $\operatorname{Traces}\left(\tau_{j}\right)$ and $t$ is in semantics of $\operatorname{AND}\left(\tau_{1}, \ldots, \tau_{j-1}, \tau_{j+1}, \ldots, \tau_{n}\right)$. Thus procedure describes from line 28 to 36 is correct.

Finally, if $\tau=\operatorname{CO}\left(\tau_{1}, \tau_{2}\right) \mathrm{n}$ then trace $t \in \operatorname{Traces}(\tau)$ if, and only if, $t \in$ $\operatorname{Traces}\left(\tau_{1}\right)$ and $t \notin \operatorname{Traces}\left(\tau_{2}\right)$, which is what is done in lines 38-45.

Proposition 5. $A D T_{k}$-NE with $k \geq 6$ is not solvable in $\operatorname{NSPACE}\left(g\left(k-5, c \sqrt{\frac{n-1}{3}}\right)\right)$.
Proof. From [21, Theorem 4.29], for an ERE $E$ of size $n$ and of $\sim$-depth $d$, there exists a constant $c$ such that the non-emptiness of $\sim E$ is not solvable in NSpace $(g(d-3, c \sqrt{n}))$. Moreover, from the proof of Theorem 2, it can be shown that the non-emptiness of $\sim E$ reduces to answering $A D T$-nE for adt $\mathrm{CO}\left(\top, \tau_{E}\right) \in A D T_{d+2}$ of size at most $3 n+1$, which concludes.

Proposition 6. For $k \geq 1, A D T_{k}-\mathrm{NE}$ is in $(k+1)$-EXPSPACE.
Proof. For an adt $\tau \in A D T_{k}$, we have a formula $\psi_{\tau} \in \Sigma_{k+1}$ with $\left|\psi_{\tau}\right| \in O\left(2^{|\tau|}\right)$ (Lemma 1) that is equivalent to $\tau$ (see Lemma 1). Now, the non-emptiness of $\operatorname{Traces}(\tau)$ is equivalent to the satisfiability of $\psi_{\tau} \in \Sigma_{k+1}$ which, by [12] take $k$-exponential time in the size of $\psi_{\tau}$ and therefore $(k+1)$-exponential time in $|\tau|$.


[^0]:    3 https://www.schneier.com/academic/archives/1999/12/attack_trees.html

[^1]:    ${ }^{4}$ arbitrary regular expressions extended with intersection and complementation.

[^2]:    ${ }^{5}$ with the convention that 0 -Expspace $=$ Pspace.

