BRICS

Basic Research in Computer Science

BRICS RS-02-47 Z. Ésik: Extended Temporal Logic on Finite Words

Extended Temporal Logic on Finite Words and Wreath Product of Monoids with Distinguished Generators

Zoltán Ésik

BRICS Report Series **RS-02-47**

ISSN 0909-0878 **December 2002**

See back inner page for a list of recent BRICS Report Series publications. Copies may be obtained by contacting:

BRICS
Department of Computer Science
University of Aarhus
Ny Munkegade, building 540
DK–8000 Aarhus C
Denmark
Telephone: +45 8942 3360
Telefax: +45 8942 3255
Internet: BRICS@brics.dk

BRICS publications are in general accessible through the World Wide Web and anonymous FTP through these URLs:

`http://www.brics.dk`
`ftp://ftp.brics.dk`
This document in subdirectory `RS/02/47/`

Extended Temporal Logic on Finite Words and Wreath Product of Monoids with Distinguished Generators

Z. Ésik*

Dept. of Computer Science, University of Szeged, P.O.B. 652, 6701 Szeged, Hungary

Abstract. We associate a modal operator with each language belonging to a given class of regular languages and use the (reverse) wreath product of monoids with distinguished generators to characterize the expressive power of the resulting logic.

1 Introduction

The wreath product and its variants have been very useful and powerful tools in the characterization of the expressive power of several logical systems over finite and infinite words, including first-order logic and its extension with modular counting, cf. Straubing [11] and Straubing, Therien, Thomas [12], temporal logic and the until hierarchy, Cohen, Pin, Perrin [3] and Therien, Thomas [13], and modular temporal logic, Bazirambawo, McKenzie, Therien [2]. In this paper, we associate a modal operator with each language in a given subclass of regular languages, and use the reverse wreath product (of monoids with distinguished generators) to provide an algebraic characterization of the expressive power of (future) temporal logic on finite words endowed with these modal operators. Our logic is closely related to that proposed in Wolper [14], and our methods are related to those of Cohen, Pin, Perrin [3] and Bazirambawo, McKenzie, Therien [2] and Ésik, Larsen [7]. Moreover, our methods and results extend to ω-words, (countable) ordinal words and, more generally, to all discrete words. These extensions will be treated in subsequent papers.

Some notation An *alphabet* is a finite nonempty set. *We assume that each alphabet is equipped with a fixed linear order* $<$. For an alphabet Σ, we denote by Σ^* the free monoid of all finite *words* over Σ including the *empty word* ϵ. The *length* of a word u is denoted $|u|$. The notation $u = u_1 \cdots u_n$ for a word $u \in \Sigma^*$ means that u is a word of length n whose letters are $u_1, \ldots, u_n$. A subset of Σ^* is called a *language* over Σ. The *boolean* and *regular operations* on languages, and *regular languages* are defined as usual. When $L \subseteq \Sigma^*$ and $u \in \Sigma^*$, the *left quotient* $u^{-1}L$ and *right quotient* Σu^{-1} of L with respect to u are respectively given by

$$u^{-1}L = \{v : uv \in L\}$$
$$Lu^{-1} = \{v : vu \in L\}.$$

* Supported in part by BRICS, a grant from the National Foundation of Hungary for Scientific Research, and by the Japan Society for the Promotion of Science.

A *class of (regular) languages* $\mathcal{L}$ consists of a set of (regular) languages for each alphabet Σ. If n is a nonnegative integer, we let $[n]$ denote the set $\{1, \ldots, n\}$. Thus, $[0]$ is another name for the empty set.

2 Extended Temporal Logic

Syntax. For an alphabet Σ, the set of formulas over Σ is the least set containing the letters p_σ, for all $\sigma \in \Sigma$, closed with respect to the boolean connectives $\vee$ (disjunction) and $\neg$ (negation), as well as the following construct. Suppose that $L \subseteq \Delta^*$ and that for each $\delta \in \Delta$, φ_δ is a formula over Σ. Then

$$L(\delta \mapsto \varphi_\delta)_{\delta \in \Delta} \tag{1}$$

is a formula over Σ. The notion of *subformula* of a formula is defined as usual.

Semantics. Suppose that φ is a formula over Σ and $u \in \Sigma^*$. We say that u *satisfies* φ, in notation $u \models \varphi$, if

- $\varphi = p_\sigma$, for some $\sigma \in \Sigma$, and u is a nonempty word whose first letter is σ, i.e., $u = \sigma u'$ for some $u' \in \Sigma^*$, or
- $\varphi = \varphi' \vee \varphi''$ and $u \models \varphi'$ or $u \models \varphi''$, or
- $\varphi = \neg\varphi'$ and it is not the case that $u \models \varphi'$, or
- $\varphi = L(\delta \mapsto \varphi_\delta)_{\delta \in \Delta}$, $u = u_1 \cdots u_n$, where each u_i is a letter, and the *characteristic word* $\delta_1 \cdots \delta_n$ belongs to L, where for each $i \in [n]$, δ_i is the first letter of Δ with respect to the linear order on Δ with $u_i \cdots u_n \models \varphi_{\delta_i}$, if such a letter exists, and δ_i is the last letter of Δ, otherwise.

For any formula φ of over Σ, we let L_φ denote the *language defined by* φ:

$$L_\varphi = \{u \in \Sigma^* : u \models \varphi\}.$$

We say that formulas φ and ψ over Σ are *equivalent* if $L_\varphi = L_\psi$. Throughout the paper we will use the boolean connective $\wedge$ (conjunction) as an abbreviation. Moreover, for any alphabet Σ, we define $\mathtt{tt} = p_\sigma \vee \neg p_\sigma$ and $\mathtt{ff} = \neg\mathtt{tt}$, where σ is a letter in Σ.

We will consider subsets of formulas associated with a class $\mathcal{L}$ of (regular) languages. We let $\mathrm{FTL}(\mathcal{L})$ denote the collection of formulas all of whose subformulas of the form (1) above satisfy that L belongs to $\mathcal{L}$. We define $\mathbf{FTL}(\mathcal{L})$ to be the class of all languages definable by formulas in $\mathrm{FTL}(\mathcal{L})$. It is clear that for each formula $L(\delta \mapsto \varphi_\delta)_{\delta \in \Delta}$ in $\mathrm{FTL}(\mathcal{L})$ over an alphabet Σ there is an equivalent formula $L(\delta \mapsto \varphi'_\delta)_{\delta \in \Delta}$ in $\mathrm{FTL}(\mathcal{L})$ such that the subformulas φ'_δ are *pairwise inconsistent*: There exists no $u \in \Sigma^*$ and distinct letters $\delta, \delta' \in \Delta$ such that $u \models \varphi'_\delta \wedge \varphi'_{\delta'}$. Indeed, when the given linear order on Δ is $\delta_1 < \cdots < \delta_k$, then we define

$$\varphi'_{\delta_i} = \varphi_{\delta_i} \wedge \bigwedge_{j<i} \neg\varphi_{\delta_j},$$

for all $i \in [k]$. Alternatively, we may define φ'_{δ_i} for all $i < k$ as above, and

$$\varphi'_{\delta_k} = \bigwedge_{j<k} \neg\varphi_{\delta_j}.$$

Thus, the modal formulas in FTL($\mathcal{L}$) over Σ associated with a language $L \subseteq \Delta^*$ in $\mathcal{L}$ may equivalently be written as $L(\delta \mapsto \varphi_\delta)_{\delta\in\Delta}$, where the φ_δ are pairwise inconsistent and $\bigvee_{\delta\in\Delta} \varphi_\delta$ is equivalent to $\mathtt{tt}$, so that each word in Σ^* satisfies exactly one φ_δ. Below we will call such families φ_δ, $\delta \in \Delta$ *deterministic*. Moreover, we will sometimes write modal formulas over Σ as $L(\delta \mapsto \varphi_\delta)_{\delta\in\Delta}$, where φ_δ, $\delta \in \Delta$ is a deterministic family of formulas over Σ. When φ_δ, $\delta \in \Delta$ is a deterministic family, we have

$$u \models L(\delta \mapsto \varphi_\delta)_{\delta\in\Delta} \Leftrightarrow \exists \delta_1 \cdots \delta_n \in L\, \forall i \in [n]\ u_i \cdots u_n \models \varphi_{\delta_i},$$

for all $u = u_1 \cdots u_n \in \Sigma^*$. We call a formula φ deterministic if for every subformula of φ of the form $L(\delta \mapsto \varphi_\delta)_{\delta\in\Delta}$, the family φ_δ, $\delta \in \Delta$ is deterministic. As shown above, for each $\varphi \in$ FTL($\mathcal{L}$) there is a deterministic formula in FTL($\mathcal{L}$) which is equivalent to φ.

We end this section with the definition of *formula substitution*. Suppose that φ is a formula over an alphabet Σ, and suppose that for each $\sigma \in \Sigma$ we are given a formula ψ_σ over Σ'. Then the formula over Σ',

$$\tau = \varphi[p_\sigma \mapsto \psi_\sigma],$$

is obtained from φ by replacing, for each letter $\sigma \in \Sigma$, each occurrence of the symbol p_σ by the formula ψ_σ. Formally, we define

- $\tau = \psi_\sigma$ if $\varphi = p_\sigma$,
- $\tau = \varphi_1[p_\sigma \mapsto \psi_\sigma] \vee \varphi_2[p_\sigma \mapsto \psi_\sigma]$ if $\varphi = \varphi_1 \vee \varphi_2$,
- $\tau = \neg(\varphi_1[p_\sigma \mapsto \psi_\sigma])$ if $\varphi = \neg\varphi_1$,
- $\tau = L(\delta \mapsto \varphi_\delta[p_\sigma \mapsto \psi_\sigma])_{\delta\in\Delta}$ if $\varphi = L(\delta \mapsto \varphi_\delta)_{\delta\in\Delta}$.

Note that when φ and ψ_σ are in FTL($\mathcal{L}$), for all $\sigma \in \Sigma$, then $\varphi[p_\sigma \mapsto \psi_\sigma]$ belongs to FTL($\mathcal{L}$).

3 Some Elementary Properties

In this section, we establish some elementary properties of the classes **FTL**($\mathcal{L}$), where $\mathcal{L}$ denotes a class of languages. We also study conditions on $\mathcal{L}$ and $\mathcal{L}'$ under which **FTL**($\mathcal{L}$) = **FTL**($\mathcal{L}'$).

Suppose that Δ and Δ' are alphabets. A *literal homomorphism* $\Delta^* \to \Delta'^*$ is a homomorphism $h : \Delta^* \to \Delta'^*$ such that $h(\Delta) \subseteq \Delta'$. Note that a homomorphism $h : \Delta^* \to \Delta'^*$ is a literal homomorphism iff it is *length preserving*, i.e., when $|h(u)| = |u|$, for all $u \in \Sigma^*$.

Proposition 1. *For each $\mathcal{L}$, the class of languages* **FTL**($\mathcal{L}$) *contains $\mathcal{L}$ and is closed with respect to the boolean operations and inverse literal homomorphisms.*

Proof. It is obvious that **FTL**($\mathcal{L}$) is closed under the boolean operations. Moreover, each language $L \subseteq \Sigma^*$ in $\mathcal{L}$ is definable by the formula $L(\sigma \mapsto p_\sigma)_{\sigma\in\Sigma}$ in FTL($\mathcal{L}$). Assume now that $h : \Sigma'^* \to \Sigma^*$ is a literal homomorphism. We argue by induction on the structure of the formula φ over Σ in FTL($\mathcal{L}$) to show that $h^{-1}(L_\varphi)$ is definable by some formula ψ in FTL($\mathcal{L}$). When $\varphi = p_\sigma$, for some letter σ, then we define $\psi = \bigvee_{h(\sigma')=\sigma} p_{\sigma'}$. It is clear that $L_\psi = h^{-1}(L_\varphi)$. Suppose now that $\varphi = \varphi_1 \vee \varphi_2$ and that $L_{\psi_i} = h^{-1}(L_{\varphi_i})$, $i = 1, 2$. Then we define $\psi = \psi_1 \vee \psi_2$. When $\varphi = \neg\varphi_1$ and $L_{\psi_1} = h^{-1}(L_{\varphi_1})$, then let $\psi = \neg\psi_1$. In either case, we have $L_\psi = h^{-1}(L_\varphi)$. Finally, assume that $\psi = L(\delta \mapsto \varphi_\delta)_{\delta\in\Delta}$, and that for each δ there is a formula ψ_δ in FTL($\mathcal{L}$) with $L_{\psi_\delta} = h^{-1}(L_{\varphi_\delta})$. Then define $\psi = L(\delta \mapsto \psi_\delta)_{\delta\in\Delta}$. Let $u = u_1 \cdots u_n \in \Sigma'^*$. Since for all $\delta \in \Delta$ and $i \in [n]$,

$$u_i \cdots u_n \models \psi_\delta \Leftrightarrow h(u_i \cdots u_n) \models \varphi_\delta,$$

the characteristic word determined by u and the formulas ψ_δ is the same as that determined by $h(u)$ and the formulas φ_δ. It follows that $u \models \psi$ iff $h(u) \models \varphi$. □

Next we show that **FTL** is a closure operator.

Proposition 2. *For any class $\mathcal{L}$ of languages,* $\mathbf{FTL}(\mathbf{FTL}(\mathcal{L})) = \mathbf{FTL}(\mathcal{L})$.

Proof. The inclusion from right to left follows from Proposition 1. To prove that $\mathbf{FTL}(\mathbf{FTL}(\mathcal{L})) \subseteq \mathbf{FTL}(\mathcal{L})$, we argue by induction on the structure of a formula φ over Δ in FTL($\mathcal{L}$) to show that for every deterministic family φ_δ, $\delta \in \Delta$ of formulas in FTL($\mathcal{L}$) over an alphabet A, the formula $L_\varphi(\delta \mapsto \varphi_\delta)_{\delta\in\Delta}$ is expressible in FTL($\mathcal{L}$), i.e., there exists a formula in FTL($\mathcal{L}$) which is equivalent to it. Assume first that $\varphi = p_{\delta_0}$, for some $\delta_0 \in \Delta$. Then $L_\varphi = \delta_0\Delta^*$. It is clear that a word $u \in A^*$ satisfies $L_\varphi(\delta \mapsto \varphi_\delta)_{\delta\in\Delta}$ iff $|u| > 0$ and u satisfies φ_{δ_0}, so that $L_\varphi(\delta \mapsto \varphi_\delta)_{\delta\in\Delta}$ is equivalent to $\varphi_{\delta_0} \wedge \bigvee_{a\in A} p_a$. In the induction step, assume first that $\varphi = \varphi_1 \vee \varphi_2$. Then $L_\varphi = L_{\varphi_1} \cup L_{\varphi_2}$ and thus $L_\varphi(\delta \mapsto \varphi_\delta)_{\delta\in\Delta}$ is equivalent to $L_{\varphi_1}(\delta \mapsto \varphi_\delta)_{\delta\in\Delta} \vee L_{\varphi_2}(\delta \mapsto \varphi_\delta)_{\delta\in\Delta}$. By induction, there exist ψ_1 and ψ_2 in FTL($\mathcal{L}$) such that $L_{\varphi_i}(\delta \mapsto \varphi_\delta)_{\delta\in\Delta}$ is equivalent to ψ_i, $i = 1, 2$. It follows that $L_\varphi(\delta \mapsto \varphi_\delta)_{\delta\in\Delta}$ is equivalent to $\psi_1 \vee \psi_2$ which is in FTL($\mathcal{L}$). Suppose next that $\varphi = \neg\varphi_1$, so that $L_\varphi = \overline{L_{\varphi_1}}$. Then we have $L_\varphi(\delta \mapsto \varphi_\delta)_{\delta\in\Delta}$ is equivalent to $\neg(L_{\varphi_1}(\delta \mapsto \varphi_\delta)_{\delta\in\Delta})$. It follows from the induction hypothesis that $L_\varphi(\delta \mapsto \varphi_\delta)_{\delta\in\Delta}$ is equivalent to a formula in FTL($\mathcal{L}$). Assume finally that $\varphi = K(\sigma \mapsto \tau_\sigma)_{\sigma\in\Sigma}$, where the family τ_σ, $\sigma \in \Sigma$ is deterministic. Then for any word $u_1 \cdots u_n \in A^*$ and for any $i \in [n]$, let φ_{δ_i} denote the unique formula φ_δ with $u_i \cdots u_n \models \varphi_\delta$. Moreover, for each $i \in [n]$, let τ_{σ_i} denote the unique formula τ_σ with $\delta_i \cdots \delta_n \models \tau_\sigma$. Then we have:

$$\begin{aligned} u_1 \cdots u_n \models L_\varphi(\delta \mapsto \varphi_\delta)_{\delta\in\Delta} &\Leftrightarrow \delta_1 \cdots \delta_n \in L_\varphi \\ &\Leftrightarrow \delta_1 \cdots \delta_n \models \varphi \\ &\Leftrightarrow \sigma_1 \cdots \sigma_n \in K. \end{aligned}$$

But for every $i \in [n]$,

$$u_i \cdots u_n \models L_{\tau_{\sigma_i}}(\delta \mapsto \varphi_\delta)_{\delta\in\Delta},$$

since for every $j \geq i$, $u_j \cdots u_n \models \varphi_{\delta_j}$ and since $\delta_i \cdots \delta_n \in L_{\tau_{\sigma_i}}$. Moreover, the formulas $L_{\tau_\sigma}(\delta \mapsto \varphi_\delta)_{\delta\in\Delta}$, $\sigma \in \Sigma$ form a deterministic family. Thus,

$$u_1 \cdots u_n \models K(\sigma \mapsto L_{\tau_\sigma}(\delta \mapsto \varphi_\delta)_{\delta\in\Delta})_{\sigma\in\Sigma} \Leftrightarrow \sigma_1 \cdots \sigma_n \in K.$$

We have thus shown that φ is equivalent to $K(\sigma \mapsto L_{\tau_\sigma}(\delta \mapsto \varphi_\delta)_{\delta \in \Delta})_{\sigma \in \Sigma}$. By the induction hypothesis, for each σ there is a formula ψ_σ in FTL($\mathcal{L}$) which is equivalent to $L_{\tau_\sigma}(\delta \mapsto \varphi_\delta)_{\delta \in \Delta}$. Thus, φ is equivalent to $K(\sigma \mapsto \psi_\sigma)_{\sigma \in \Sigma}$. □

Since $\mathbf{FTL}(\mathcal{L}_1) \subseteq \mathbf{FTL}(\mathcal{L}_2)$ whenever $\mathcal{L}_1 \subseteq \mathcal{L}_2$ and since $\mathcal{L} \subseteq \mathbf{FTL}(\mathcal{L})$, for all $\mathcal{L}$, we have:

Corollary 1. **FTL** *is a closure operator.*

Proposition 3. *Suppose that for each formula $L(\delta \mapsto \varphi_\delta)_{\delta \in \Delta}$ in* FTL($\mathcal{L}$)*, over any alphabet Σ, and for each $w \in \Delta^*$ there is a formula in* FTL($\mathcal{L}$) *which is equivalent to $(w^{-1}L)(\delta \mapsto \varphi_\delta)_{\delta \in \Delta}$. Then* $\mathbf{FTL}(\mathcal{L})$ *is closed with respect to left quotients.*

Proof. Suppose that φ is a formula over Σ in FTL($\mathcal{L}$) and σ is a letter in Σ. We show that $\sigma^{-1}L_\varphi$ belongs to $\mathbf{FTL}(\mathcal{L})$. The generalization to quotients $w^{-1}L_\varphi$, where w is a word, is left to the reader. When φ is p_σ, then $\sigma^{-1}L_\varphi$ is Σ^*, which is definable by the formula $\mathtt{tt}$. When φ is $p_{\sigma'}$, where $\sigma' \neq \sigma$, then $\sigma^{-1}L_\varphi$ is $\emptyset$, which is definable by the formula $\mathtt{ff}$. We continue by induction on the structure of φ. Suppose that $\varphi = \varphi_1 \vee \varphi_2$ or $\varphi = \neg\varphi_1$, and assume that $\sigma^{-1}L_{\varphi_i}$ is defined by $\widetilde{\varphi}_i$ in FTL($\mathcal{L}$), $i = 1, 2$. Then $\sigma^{-1}L_\varphi$ is defined by $\widetilde{\varphi}_1 \vee \widetilde{\varphi}_2$ or $\neg\widetilde{\varphi}_1$, respectively. Assume finally that φ is $L(\delta \mapsto \varphi_\delta)_{\delta \in \Delta}$, where φ_δ, $\delta \in \Delta$ is a deterministic family, and that for each δ, $\sigma^{-1}L_{\varphi_\delta}$ is defined by $\widetilde{\varphi}_\delta$ in FTL($\mathcal{L}$). Note that $\widetilde{\varphi}_\delta$, $\delta \in \Delta$ is also a deterministic family. By assumption, for each δ_0 in Δ there is a formula τ_{δ_0} in FTL($\mathcal{L}$) such that for all words $u \in \Sigma^*$,

$$u \models \tau_{\delta_0} \Leftrightarrow u \models (\delta_0^{-1}L)(\delta \mapsto \varphi_\delta)_{\delta \in \Delta}.$$

Then let

$$\widetilde{\varphi} = \bigvee_{\delta_0 \in \Delta} (\widetilde{\varphi}_{\delta_0} \wedge \tau_{\delta_0}).$$

We have, for all $u = u_1 \cdots u_n \in \Sigma^*$,

$$\begin{aligned}
u \models \widetilde{\varphi} &\Leftrightarrow \exists \delta_0 \; u \models \widetilde{\varphi}_{\delta_0} \wedge u \models \tau_{\delta_0} \\
&\Leftrightarrow \exists \delta_0 \; \sigma u \models \varphi_{\delta_0} \wedge u \models (\delta_0^{-1}L)(\delta \mapsto \varphi_\delta)_{\delta \in \Delta} \\
&\Leftrightarrow \exists \delta_0 \; \sigma u \models \varphi_{\delta_0} \wedge \exists \delta_1 \cdots \delta_n \in \delta_0^{-1}L \; \forall i \in [n] \; u_i \cdots u_n \models \varphi_{\delta_i} \\
&\Leftrightarrow \exists \delta_0 \cdots \delta_n \in L \; \sigma u \models \varphi_{\delta_0} \wedge \forall i \in [n] \; u_i \cdots u_n \models \varphi_{\delta_i} \\
&\Leftrightarrow \sigma u \models L(\delta \mapsto \varphi_\delta)_{\delta \in \Delta} \\
&\Leftrightarrow \sigma u \models \varphi.
\end{aligned}$$

This concludes the proof of Proposition 3. □

Proposition 4. *Suppose that for each formula $L(\delta \mapsto \varphi_\delta)_{\delta \in \Delta}$ in* FTL($\mathcal{L}$)*, over any alphabet Σ, and for each $w \in \Delta^*$ there is a formula in* FTL($\mathcal{L}$) *which is equivalent to $(Lw^{-1})(\delta \mapsto \varphi_\delta)_{\delta \in \Delta}$. Then* $\mathbf{FTL}(\mathcal{L})$ *is closed with respect to right quotients.*

Proof. Suppose that φ is a formula over Σ in FTL($\mathcal{L}$), and let σ be a letter in Σ. We only show that $L_\varphi \sigma^{-1}$ belongs to $\mathbf{FTL}(\mathcal{L})$. When φ is p_σ, then $L_\varphi \sigma^{-1}$ is $\sigma\Sigma^* \cup \{\epsilon\}$, which is defined by the formula $p_\sigma \vee \bigwedge_{\sigma' \in \Sigma} \neg p_{\sigma'}$. When φ is $p_{\sigma'}$, where $\sigma' \neq \sigma$, then $\sigma^{-1} L_\varphi$ is $\sigma' \Sigma^*$, which is defined by the formula $p_{\sigma'}$. We proceed by induction on the structure of φ. The cases when $\varphi = \varphi_1 \vee \varphi_2$ or $\varphi = \neg\varphi_1$ can be handled as above. Assume finally that φ is $L(\delta \mapsto \varphi_\delta)_{\delta \in \Delta}$, and that for each δ, $L_{\varphi_\delta}\sigma^{-1}$ is defined by $\widetilde{\varphi}_\delta$ in FTL($\mathcal{L}$). By assumption, for each δ_0 in Δ there is a formula τ_{δ_0} such that for all words $u \in \Sigma^*$,

$$u \models \tau_{\delta_0} \Leftrightarrow u \models (L\delta_0^{-1})(\delta \mapsto \widetilde{\varphi}_\delta)_{\delta \in \Delta}.$$

We define

$$\widetilde{\varphi} = \bigvee_{\sigma \models \varphi_{\delta_0}} \tau_{\delta_0}.$$

Then, $u \in L_\varphi \sigma^{-1}$ iff $u\sigma \in L_\varphi$ iff the characteristic word determined by $u\sigma$ and the formulas φ_δ belongs to L iff there exists some δ_0 such that $\sigma \models \varphi_{\delta_0}$ and the characteristic word determined by u and the formulas $\widetilde{\varphi}_\delta$ belongs to $L\delta_0^{-1}$ iff $u \models \widetilde{\varphi}$. □

Corollary 2. *1. For any class $\mathcal{L}$ of languages,* $\mathbf{FTL}(\mathcal{L}) = \mathbf{FTL}(\mathcal{L}')$, *where $\mathcal{L}'$ is the least class containing $\mathcal{L}$ closed with respect to the boolean operations and inverse literal morphisms.*

2. For any class $\mathcal{L}$ of languages closed with respect to quotients, or such that the modal operators associated with the quotients of the languages in $\mathcal{L}$ are expressible in FTL($\mathcal{L}$) *as in Propositions 3 and 4,* $\mathbf{FTL}(\mathcal{L}) = \mathbf{FTL}(\mathcal{L}')$, *where $\mathcal{L}'$ is the least class containing $\mathcal{L}$ closed with respect to the boolean operations, quotients, and inverse literal morphisms.*

4 Monoids with Distinguished Generators

Suppose that M is a monoid and A is a nonempty set of distinguished generators for M. Then we call the pair (M, A) a *monoid with distinguished generators*, or *mg-pair*, for short. When M is finite, the mg-pair (M, A) is also called finite.

Suppose that (M, A) and (N, B) are mg-pairs. A *homomorphism* $(M, A) \to (N, B)$ is a monoid homomorphism $h : M \to N$ such that $h(A) \subseteq B$. It is clear that mg-pairs equipped with these homomorphisms form a category. We call (M, A) a *sub mg-pair* of (N, B) if M is a submonoid of N and A is a subset of B. Moreover, we call (M, A) a *quotient* of (N, B) if there is a surjective homomorphism $(N, B) \to (M, A)$, i.e., a homomorphism of mg-pairs which maps B onto A. We say that (M, A) *divides*, or *is a divisor of* (N, B), denoted $(M, A) < (N, B)$, if (M, A) is a quotient of a sub mg-pair of (N, B). *We identify any monoid M with the mg-pair (M, M).*

Example 1. For every alphabet Σ, (Σ^*, Σ) is an mg-pair with the following property: For every mg-pair (M, A) and function $h : \Sigma \to A$ there is a unique homomorphism $h^\sharp : (\Sigma^*, \Sigma) \to (M, A)$ extending h. We call such mg-pairs *free.*

Let $L \subseteq \Sigma^*$. Recall that the *syntactic monoid* of L is the quotient $\Sigma^*/\sim_L$ of Σ^* with respect to the syntactic congruence $\sim_L$ defined on Σ^* by

$$u \sim_L v \Leftrightarrow \forall x, y \in \Sigma^* \; xuy \in L \Leftrightarrow xvy \in L.$$

The *syntactic mg-pair* of L is $Synt(L) = (\Sigma^*/\sim_L, \Sigma/\sim_L)$.

We call a language $L \subseteq \Sigma^*$ *recognizable by an mg-pair* (M, A) if there is a homomorphism $h : (\Sigma^*, \Sigma) \to (M, A)$ with $L = h^{-1}(h(L))$. It follows by standard arguments that a language L is recognizable by an mg-pair (M, A) iff $Synt(L) < (M, A)$. Moreover, a language is recognizable by a finite mg-pair iff it is regular.

For the definition of the *(reverse) semidirect product* of monoids we refer to Eilenberg [5], and for the extension of these notions to mg-pairs to Ésik and Larsen [7]. When (S, A) and (T, B) are mg-pairs equipped with a (monoidal) *right action* of T on S,

$$\begin{aligned} S \times T &\to S \\ (s, t) &\mapsto st \end{aligned}$$

such that $st \in A$ whenever $s \in A$, we let $(S, A) \star_r (T, B)$ denote the reverse semidirect product of (S, A) and (T, B) determined by the right action. This is the mg-pair $(R, A \times B)$, where R is the submonoid of the ordinary reverse semidirect product $S \star_r T$ of the monoids S and T determined by the action. When the right action is trivial, i.e., $st = s$ for all $s \in S$ and $t \in T$, the reverse semidirect product $(S, A) \star_r (T, B)$ becomes the *direct product* $(S, A) \times (T, B)$, i.e., the mg-pair $(R, A \times B)$, where R is the submonoid generated by $A \times B$ in the usual direct product $S \times T$ of the monoids S and T.

In addition to the reverse semidirect product, we will also make use of the *reverse wreath product*. Suppose that (S, A) and (T, B) are mg-pairs. Then consider the direct power $(S, A)^T$ of (S, A), i.e., the mg-pair (R, A^T), where R is the submonoid of S^T generated by A^T. Define the right action of T on R by

$$(ft)(t') = f(tt'),$$

for all $f \in R$ and $t, t' \in T$. Then the reverse wreath product $(S, A) \circ_r (T, B)$ is the reverse semidirect product $(R, A^T) \star_r (T, B)$ determined by the above action.

In the sequel, except for free mg-pairs, we will only consider finite mg-pairs. We call a nonempty class of finite mg-pairs a *variety* if it is closed with respect to division and direct product. A *closed variety* is also closed with respect the reverse semidirect product (or reverse wreath product). For any class $\mathbf{K}$ of finite mg-pairs, we let $\widehat{\mathbf{K}}$ denote the least closed variety containing $\mathbf{K}$. An example of a closed variety is the class $\mathbf{D}^r$ of all *reverse definite mg-pairs*. We call a finite mg-pair (M, A) reverse definite if there exists an integer $n \geq 0$ such that $a_1 \cdots a_n = a_1 \cdots a_{n+1}$ for all $a_1, \ldots, a_{n+1}$ in A. For example, when M_n denotes the monoid of all words over the two-letter alphabet $\{a, b\}$ whose length is at most n equipped with the product operation $u \cdot v = w$ iff w is the maximal prefix of uv of length $\leq n$, then $E_n = (M_n, \{a, b\})$ is a reverse definite mg-pair. Below we will write E for E_2. Note that E_n is a quotient of E_{n+1}, for all n. Each E_n generates $\mathbf{D}^r$:

Proposition 5. *For each $n \geq 1$, $\mathbf{D}^r$ is the least closed variety containing E_n.*

This follows by adapting the proof of a well-known fact for definite semigroups, proved in Eilenberg [5]. Further examples of closed varieties will be introduced when needed. When $\mathbf{V}$ and $\mathbf{W}$ are closed varieties, we let $\mathbf{V} \vee \mathbf{W}$ denote the least *closed* variety containing $\mathbf{V} \cup \mathbf{W}$.

Suppose that $\mathbf{K}$ is a class of finite mg-pairs. We let $\mathcal{L}_{\mathbf{K}}$ denote the class of all regular languages recognizable by the mg-pairs in $\mathbf{K}$. By standard arguments, it follows that a language is in $\mathcal{L}_{\mathbf{K}}$ iff its syntactic mg-pair is in the variety generated by $\mathbf{K}$. Conversely, when $\mathcal{L}$ is a class of regular languages, let $\mathbf{K}_{\mathcal{L}}$ denote the class of all syntactic mg-pairs of the languages in $\mathcal{L}$.

For each class $\mathbf{K}$ of finite mg-pairs, we define $\mathrm{FTL}(\mathbf{K}) = \mathrm{FTL}(\mathcal{L}_{\mathbf{K}})$ and $\mathbf{FTL}(\mathbf{K}) = \mathbf{FTL}(\mathcal{L}_{\mathbf{K}})$.

Corollary 3. *Let $\mathcal{L}$ denote a class of regular languages. We have $\mathbf{FTL}(\mathcal{L}) = \mathbf{FTL}(\mathbf{K}_{\mathcal{L}})$ iff there exists some class $\mathbf{K}$ of finite mg-pairs with $\mathbf{FTL}(\mathcal{L}) = \mathbf{FTL}(\mathbf{K})$ iff for each $L \subseteq \Delta^*$ in $\mathcal{L}$ and for each $w \in \Delta^*$, the modal operators associated with $w^{-1}L$ and Lw^{-1} are expressible in* $\mathrm{FTL}(\mathcal{L})$ *as in Propositions 3 and 4.*

Remark 1. Given a class $\mathbf{K}$ of finite mg-pairs, let $\mathcal{L}$ denote the class of all regular languages $L \subseteq A^*$ such that there exists some mg-pair $(S, A) \in \mathbf{K}$ with $L = h^{-1}(h(L))$, where h denotes the homomorphism $(A^*, A) \to (S, A)$ which is the identity function on A. Then it follows from Corollary 2 and Corollary 1 that $\mathbf{FTL}(\mathbf{K}) = \mathbf{FTL}(\mathcal{L})$.

5 Main Results

We say that the *next modality is expressible in* $\mathrm{FTL}(\mathcal{L})$ if for each formula φ in $\mathrm{FTL}(\mathcal{L})$ over any alphabet Σ there is a formula $\mathsf{X}\varphi$ over Σ such that for all $u \in \Sigma^*$,

$$u \models \mathsf{X}\varphi \Leftrightarrow \exists \sigma \in \Sigma, v \in \Sigma^* \; u = \sigma v \wedge v \models \varphi.$$

Proposition 6. *The next modality is expressible in* $\mathrm{FTL}(\mathcal{L})$ *iff the two-letter regular language $(a+b)b(a+b)^*$ and the one-letter language a belong to* $\mathbf{FTL}(\mathcal{L})$.

Proof. Suppose first that $L_1 = (a+b)b(a+b)^*$ and $L_2 = a$ are in $\mathbf{FTL}(\mathcal{L})$. Let φ be any formula in $\mathrm{FTL}(\mathcal{L})$ over the alphabet Σ. If $\epsilon \not\models \varphi$, then $\mathsf{X}\varphi$ is expressible as $L_1(a \mapsto \neg\varphi, b \mapsto \varphi)$. If $\epsilon \models \varphi$, then $X\varphi$ is expressible by $L_1(a \mapsto \neg\varphi, b \mapsto \varphi) \vee L_2(a \mapsto \mathtt{tt})$. It follows from Corollary 1 that the next modality is expressible in $\mathrm{FTL}(\mathcal{L})$.

Suppose now that the next modality is expressible in $\mathrm{FTL}(\mathcal{L})$, so that $\mathsf{X}\varphi$ exists for each formula φ in FTL. Then $(a+b)b(a+b)^*$ is definable by $\mathsf{X}p_b$, and a is definable by $p_a \wedge \mathsf{X}\neg p_a$. □

Note that the mg-pair E defined above is isomorphic to the syntactic mg-pair of the language $(a+b)b(a+b)^*$. Using this fact, we have:

Corollary 4. *For any class* $\mathbf{K}$ *of finite mg-pairs, the next modality is expressible in* $\mathrm{FTL}(\mathbf{K})$ *iff* $(a+b)b(a+b)^*$ *belongs to* $\mathbf{FTL}(\mathbf{K})$ *iff every language recognizable by* E *belongs to* $\mathbf{FTL}(\mathbf{K})$.

Proof. If $L = (a+b)b(a+b)^*$ belongs to $\mathbf{FTL}(\mathbf{K})$, then, by Corollary 2, so does a, regarded as a one-letter language, since it can be constructed from L by the boolean operations, left quotients, and inverse literal homomorphisms. The first equivalence in Corollary 4 now follows from Proposition 6. As for the second, it is clear that if every language recognizable by E belongs to $\mathbf{FTL}(\mathbf{K})$, then so does the language $(a+b)b(a+b)^*$, since it is recognizable by E. On the other hand, it can be shown by standard arguments that any language recognizable by the syntactic mg-pair of a regular language L is the inverse image under a literal homomorphism of a boolean combination of quotients of L. Thus, since E is the syntactic monoid of $(a+b)b(a+b)^*$, if $(a+b)b(a+b)^*$ is in $\mathbf{FTL}(\mathbf{K})$, then since $\mathbf{FTL}(\mathbf{K})$ is closed with respect to the above operations, it follows that every language recognizable by E is in $\mathbf{FTL}(\mathbf{K})$. □

Corollary 5. *For any class* $\mathbf{K}$ *of finite mg-pairs, the next modality is expressible in* $\mathrm{FTL}(\mathbf{K})$ *iff* $\mathbf{FTL}(\mathbf{K}) = \mathbf{FTL}(\mathbf{K}_1)$*, where* $\mathbf{K}_1 = \mathbf{K} \cup \{E\}$.

Proof. By Corollary 4 and Corollary 1. □

Proposition 7. *Suppose that* (S, A) *and* (T, B) *are finite mg-pairs and* $(R, A \times B)$ *is a reverse semidirect product of* (S, A) *and* (T, B) *determined by a right action of* T *on* S. *If every language recognizable by* (S, A) *and* (T, B) *belongs to* $\mathbf{FTL}(\mathcal{L})$*, and if the next modality is expressible in* $\mathrm{FTL}(\mathcal{L})$*, then every language recognizable by* $(R, A \times B)$ *also belongs to* $\mathbf{FTL}(\mathcal{L})$.

Proof. Let h denote a homomorphism

$$
\begin{aligned}
(\Sigma^*, \Sigma) &\to (R, A \times B) \\
u &\mapsto h(u) = (h_\ell(u), h_r(u)).
\end{aligned}
$$

It suffices to show that for each $(s, t) \in R$, the language $h^{-1}((s, t))$ belongs to $\mathbf{FTL}(\mathcal{L})$.

For each $u \in A^*$, let $\overline{u}$ denote the image of u under the homomorphism $(A^*, A) \to (S, A)$ which is the identity map on A. Moreover, let $\hat{h}$ denote the function $\Sigma^* \to A^*$ defined by

$$
\begin{aligned}
\hat{h}(\epsilon) &= \epsilon \\
\hat{h}(\sigma u) &= [h_\ell(\sigma) h_r(u)]\hat{h}(u),
\end{aligned}
$$

for all $u \in \Sigma^*$ and $\sigma \in \Sigma$. Here, $h_\ell(\sigma)h_r(u)$ is the result of the right action of $h_r(u)$ on $h_\ell(\sigma)$. Note that $|\hat{h}(u)| = |u|$ and that $\hat{h}(u)$ is a suffix of $\hat{h}(v)$ whenever u is a suffix of v. Also, $\overline{\hat{h}(u)} = h_\ell(u)$, for all $u \in \Sigma^*$.

By assumption, for each $s \in S$ and $t \in T$ there exist a formula φ_s over A and a formula φ_t over Σ in FTL($\mathcal{L}$) such that

$$\begin{aligned} L_{\varphi_s} &= \{w \in A^* : \overline{w} = s\} \\ L_{\varphi_t} &= \{u \in \Sigma^* : h_r(u) = t\}. \end{aligned}$$

Given a formula ψ over A in FTL($\mathcal{L}$), let

$$\psi' = \psi[p_a \mapsto \bigvee_{h_\ell(\sigma)t=a} (p_\sigma \wedge \mathsf{X}\varphi_t)].$$

Claim For all $u \in \Sigma^*$,

$$u \models \psi' \Leftrightarrow \hat{h}(u) \models \psi.$$

We prove this claim by induction on the structure of ψ. When ψ is p_a, for some $a \in A$, we have

$$\begin{aligned} u \models \psi' &\Leftrightarrow \exists \sigma, t[u \models p_\sigma \wedge u \models \mathsf{X}\varphi_t \wedge h_\ell(\sigma)t = a] \\ &\Leftrightarrow \exists \sigma, t, v[u = \sigma v \wedge v \models \varphi_t \wedge h_\ell(\sigma)t = a] \\ &\Leftrightarrow \exists \sigma, t, v[u = \sigma v \wedge h_r(v) = t \wedge h_\ell(\sigma)t = a] \\ &\Leftrightarrow \exists \sigma, v[u = \sigma v \wedge h_\ell(\sigma)h_r(v) = a] \\ &\Leftrightarrow \hat{h}(u) \models p_a \\ &\Leftrightarrow \hat{h}(u) \models \psi. \end{aligned}$$

The induction step is obvious when ψ is the disjunction $\psi_1 \vee \psi_2$ or a negation $\neg\psi_1$. Suppose now that ψ is $L(\delta \mapsto \psi_\delta)_{\delta \in \Delta}$, where $L \subseteq \Delta^*$ is in $\mathcal{L}$ and each φ_δ is a formula in FTL($\mathcal{L}$) over Σ satisfying the induction assumption. Suppose that $u = u_1 \cdots u_n$, say. By the induction hypothesis we have that for all $i \in [n]$ and $\delta \in \Delta$,

$$u_i \cdots u_n \models \psi'_\delta \Leftrightarrow \hat{h}(u_i \cdots u_n) \models \psi_\delta.$$

Thus, since $\hat{h}$ preserves suffixes, the characteristic word determined by u and ψ' is the same as that determined by $\hat{h}(u)$ and ψ, proving that $u \models \psi'$ iff $\hat{h}(u) \models \psi$.

We now complete the proof of the proposition. For any $(s,t) \in R$ and $u \in \Sigma^*$,

$$\begin{aligned} h(u) = (s,t) &\Leftrightarrow h_\ell(u) = s \wedge h_r(u) = t \\ &\Leftrightarrow \overline{\hat{h}(u)} = s \wedge h_r(u) = t \\ &\Leftrightarrow \hat{h}(u) \models \varphi_s \wedge u \models \varphi_t \\ &\Leftrightarrow u \models \varphi'_s \wedge \varphi_t. \end{aligned}$$

Thus, $\varphi'_s \wedge \varphi_t$ defines $h^{-1}((s,t))$. □

Proposition 8. *Suppose that* $\varphi = K(\delta \mapsto \varphi_\delta)_{\delta\in\Delta}$ *is a formula over* Σ *in* FTL($\mathcal{L}$)*, where* φ_δ*,* $\delta \in \Delta$ *is a deterministic family. Suppose that* K *is recognized by* $h_K : (\Delta^*, \Delta) \to (S, A)$ *and that each* L_{φ_δ} *is recognized by the morphism* $h : (\Sigma^*, \Sigma) \to (T, B)$*. Then* $L_\varphi \subseteq \Sigma^*$ *is recognizable by the reverse wreath product* $(S, A) \circ_r (T, B)$*.*

Proof. Without loss of generality we may assume that h is surjective. For each $\delta \in \Delta$, let F_δ denote the set $h(L_{\varphi_\delta})$. Note that the sets F_δ are pairwise disjoint by assumption, and that $\bigcup_{\delta\in\Delta} F_\delta = T$, since h is surjective. Define $k : (\Sigma^*, \Sigma) \to (S, A) \circ_r (T, B)$ by

$$k(\sigma) = (f_\sigma, h(\sigma)),$$

for all $\sigma \in \Sigma$, where for each $t \in T$, $f_\sigma(t) = h_K(\delta)$ for the unique δ with $h(\sigma)t \in F_\delta$. Let $u = u_1 \cdots u_n$ in Σ^*. We have

$$k(u) = (f, h(u)),$$

where

$$f(t) = h_K(\delta_1 \cdots \delta_n)$$

for the unique word $\delta_1 \cdots \delta_n$ with $h(u_i \cdots u_n)t \in F_{\delta_i}$ for each $i \in [n]$. In particular,

$$f(1) = h_K(\delta_1 \cdots \delta_n)$$

for the unique word $\delta_1 \cdots \delta_n$ with $h(u_i \cdots u_n) \in F_{\delta_i}$, i.e.,

$$u_i \cdots u_n \models \varphi_{\delta_i}$$

for each $i \in [n]$. Thus,

$$\begin{aligned} f(1) \in h_K(K) &\Leftrightarrow \delta_1 \cdots \delta_n \in K \\ &\Leftrightarrow u \models \varphi. \end{aligned}$$

It follows that

$$L_\varphi = \{u \in \Sigma^* : f(1) \in h_K(K)\}.$$

This proves that L_φ is recognizable by $(S, A) \circ_r (T, B)$. □

Recall that E_1 denotes the mg-pair $(\{1, a, b\}, \{a, b\})$, where 1 is the identity element and a, b are both left-zeroes. Note that E_1 is just the syntactic mg-pair of the language $a(a + b)^*$.

Theorem 1. *For any class* **K** *of finite mg-pairs, every language in* **FTL**(**K**) *is recognizable by some mg-pair in* $\widehat{\mathbf{K}} \vee \mathbf{D}^r$*.*

Proof. Let φ denote a deterministic formula over Σ in FTL($\mathbf{K}$). When φ is p_σ, for some $\sigma \in \Sigma$, then $L_\varphi = \sigma\Sigma^*$, which is recognizable by $E_1 \in \mathbf{D}^r$. We continue by induction on the structure of φ. Assume that $\varphi = \varphi_1 \vee \varphi_2$ such that L_{φ_i} is recognizable by (S_i, A_i) in $\widehat{\mathbf{K}} \vee \mathbf{D}^r$, $i = 1, 2$. Then L_φ is recognizable by the direct product $(S_1, A_1) \times (S_2, A_2)$ which is also in $\widehat{\mathbf{K}} \vee \mathbf{D}^r$. When $\varphi = \neg\varphi_1$, where L_{φ_1} is recognizable by (S_1, A_1) above, then L_φ is recognizable by the same mg-pair (S_1, A_1). Finally, when $\varphi = L(\delta \mapsto \varphi_\delta)_{\delta\in\Delta}$ and each L_{φ_δ} is recognizable by some mg-pair in $\widehat{\mathbf{K}} \vee \mathbf{D}^r$, then it follows by Proposition 8 that L_φ is recognizable by some mg-pair in $\widehat{\mathbf{K}} \vee \mathbf{D}^r$. (Note that since $\widehat{\mathbf{K}} \vee \mathbf{D}^r$ is closed with respect to the direct product, we may assume without loss of generality that each L_{φ_δ} is recognizable by the same mg-pair (M, A) in $\widehat{\mathbf{K}} \vee \mathbf{D}^r$, and in fact by the same morphism $(\Sigma^*, \Sigma) \to (M, A)$.) □

Theorem 2. *Suppose that the next modality is expressible in* FTL($\mathbf{K}$)*, where* $\mathbf{K}$ *is a class of finite mg-pairs. Then a language L belongs to* $\mathbf{FTL}(\mathbf{K})$ *iff* $Synt(L)$ *belongs to* $\widehat{\mathbf{K}} \vee \mathbf{D}^r$.

Proof. First, by Corollary 5, we have $\mathbf{FTL}(\mathbf{K}) = \mathbf{FTL}(\mathbf{K}_1)$, where $\mathbf{K}_1 = \mathbf{K} \cup \{E\}$, so that $\widehat{\mathbf{K}}_1 = \widehat{\mathbf{K}} \vee \mathbf{D}^r$. Let us define the *rank* of $(S, A) \in \widehat{\mathbf{K}} \vee \mathbf{D}^r$ to be the smallest number of reverse semidirect product and division operations needed to generate (S, A) from $\mathbf{K}_1$. We prove by induction on the rank of (S, A) that every language recognizable by (S, A) is in $\mathbf{FTL}(\mathbf{K}_1)$. When the rank is 0 we have $(S, A) \in \mathbf{K}_1$. Thus the result follows from Proposition 1. When the rank of (S, A) is positive, then (S, A) either divides an mg-pair (T, B) in $\widehat{\mathbf{K}} \vee \mathbf{D}^r$ of smaller rank, or (S, A) is the reverse semidirect product of some mg-pairs in $\widehat{\mathbf{K}} \vee \mathbf{D}^r$ of smaller rank. In the first case, every language recognizable by (S, A) is recognizable by (T, B). In the second case, the result follows from Proposition 7.

To necessity part of Theorem 2 follows from Theorem 1. □

Corollary 6. *For each class $\mathcal{L}$ of regular languages,* $\mathbf{FTL}(\mathcal{L})$ *consists of regular languages.*

Call a nonempty class of regular languages $\mathcal{L}$ *closed* if $\mathbf{FTL}(\mathcal{L}) \subseteq \mathcal{L}$ and if $\mathcal{L}$ is closed with respect to quotients. By Propositions 1, 3 and 4, every closed class is a *literal variety*, i.e., it is closed with respect to the boolean operations, quotients, and inverse literal homomorphisms. Moreover, by Corollaries 1, 2 and 3, $\mathcal{L}$ is closed iff $\mathcal{L} = \mathbf{FTL}(\mathcal{L}')$ for a class $\mathcal{L}'$ of regular languages closed with respect to quotients iff $\mathcal{L} = \mathbf{FTL}(\mathbf{K})$ for a class $\mathbf{K}$ of finite mg-pairs.

The assignment $\mathbf{V} \mapsto \mathcal{L}_\mathbf{V}$ defines an order isomorphism between varieties $\mathbf{V}$ of finite mg-pairs and literal varieties of regular languages, cf. Ésik, Larsen [7]. The inverse assignment maps a literal variety $\mathcal{L}$ to the class of those finite mg-pairs (M, A) such that every language recognizable by (M, A) belongs to $\mathcal{L}$.

Theorem 3. *The assignment* $\mathbf{V} \mapsto \mathcal{L}_\mathbf{V} = \mathbf{FTL}(\mathbf{V})$ *defines an order isomorphism between closed varieties* $\mathbf{V}$ *of finite mg-pairs containing* $\mathbf{D}^r$ *and closed classes* $\mathcal{L}$ *of regular languages containing* $(a + b)b(a + b)^*$.

Proof. If $\mathbf{V}$ is a closed variety containing $\mathbf{D}^r$, then by Theorem 1, $\mathcal{L}_{\mathbf{V}}$ is a closed class of regular languages containing $(a+b)b(a+b)^*$. Since $\mathbf{FTL}(\mathbf{V})$ is the least closed class containing $\mathcal{L}_{\mathbf{V}}$, it follows that $\mathcal{L}_{\mathbf{V}} = \mathbf{FTL}(\mathbf{V})$. As mentioned above, we have $\mathbf{V}_1 \subseteq \mathbf{V}_2$ iff $\mathcal{L}_{\mathbf{V}_1} \subseteq \mathcal{L}_{\mathbf{V}_2}$. Finally, the map is surjective, for if $\mathcal{L}$ is a closed class of regular languages containing $(a+b)b(a+b)^*$, then $\mathcal{L} = \mathcal{L}_{\mathbf{V}}$ for some variety $\mathbf{V}$ of finite mg-pairs containing E. By Proposition 7, $\mathbf{V}$ is closed with respect to the reverse semidirect product. Since $\mathbf{V}$ contains E, by Proposition 5 it also contains all the reverse definite mg-pairs. □

We refer to Almeida [1] for a detailed study of varieties of finite semigroups closed with respect to the semidirect product. Any such variety gives rise to a closed variety of finite mg-pairs.

Example 2. The closed class of regular languages corresponding to $\mathbf{D}^r$ is the class of reverse definite languages, where a language $L \subseteq \Sigma^*$ is termed reverse definite iff there is some $n \geq 0$ such that the membership of a word u in L depends only on the maximal prefix of u of length $\leq n$. (This condition is equivalent to requiring that the language is recognizable by some E_n.)

6 Some Applications

Propositional (future) temporal logic (FTL) was introduced in Pnueli [10]. The formulas of FTL over an alphabet Σ are constructed from the letters p_σ, $\sigma \in \Sigma$ by the boolean connectives $\vee$ and $\neg$ and the *next* and *until* modalities, denoted $\mathsf{X}\varphi$ and $\varphi\mathsf{U}\psi$. The semantics of FTL are defined similarly to that of $\mathbf{FTL}(\mathcal{L})$. In particular, when $u = u_1 \cdots u_n \in \Sigma^*$ and φ and ψ are formulas over Σ,

1. $u \models \mathsf{X}\varphi$ iff $n \geq 1$ and $u_2 \cdots u_n \models \varphi$,
2. $u \models \varphi\mathsf{U}\psi$ iff there exists some $i \in [n]$ such that $u_i \cdots u_n \models \psi$ and $u_j \cdots u_n \models \varphi$ for all $j < i$.

We let $\mathbf{FTL}$ denote the class of languages definable by the formulas in FTL.

Let U^r denote the monoid component of E_1, i.e., the monoid $\{1, a, b\}$, where a, b are left-zero elements.

Proposition 9. $\mathbf{FTL} = \mathbf{FTL}(\{U^r, E\})$.

Proof. The inclusion $\mathbf{FTL} \subseteq \mathbf{FTL}(\{U^r, E\})$ follows from Corollary 4 and the fact that $\varphi\mathsf{U}\psi$ is expressible as $L_{\mathsf{U}}(a \mapsto \varphi \wedge \neg\psi,\ b \mapsto \psi,\ c \mapsto \neg\varphi \vee \neg\psi)$, where L_{U} denotes the language $a^*b(a+b+c)^*$ over the three-letter alphabet $\{a, b, c\}$ which can be recognized by U^r. For the reverse inclusion, one can show easily that every language recognizable by U^r or E is definable in $\mathbf{FTL}$. For example, L_{U} is defined by $p_a\mathsf{U}p_b$. It then follows that the modal operator corresponding to any language recognized by U^r or E is expressible in $\mathbf{FTL}$. □

We call a finite mg-pair *aperiodic* if its monoid component is aperiodic, cf. Eilenberg [5]. For example, U^r and E are aperiodic. It follows from (the dual of) the Krohn-Rhodes Decomposition Theorem [5] that the aperiodic mg-pairs form a closed variety, namely the closed variety generated by U^r. Let $\mathbf{A}$ denote the closed variety of aperiodic mg-pairs. For *first-order definable* languages, we refer to [9].

Theorem 4. Mc Naughton-Papert [9] *A language is first-order definable iff* $Synt(L) \in \mathbf{A}$.

Theorem 5. *Let* $\mathbf{K}$ *denote a class of finite mg-pairs such that the next modality is expressible in* $\mathrm{FTL}(\mathbf{K})$. *Then* $\mathbf{FTL}(\mathbf{K})$ *contains the first-order definable languages iff* $\mathbf{A} \subseteq \widehat{\mathbf{K}} \vee \mathbf{D}^r$. *Moreover,* $\mathbf{FTL}(\mathbf{K})$ *is the class of all first-order definable languages iff* $\mathbf{A} = \widehat{\mathbf{K}} \vee \mathbf{D}^r$.

Proof. Immediate from Theorem 2, Theorem 3, the Krohn-Rhodes Decomposition Theorem and the fact that E is aperiodic. □

From Theorem 5 and Proposition 9 we immediately have:

Corollary 7. Kamp [8] *A language is first-order definable iff it is in* $\mathbf{FTL}$.

We now consider cyclic counting. For any formula φ over Σ and integers $k \geq 1$ and $0 \leq r < k$, let $C_k^r\varphi$ be a formula with the following semantics. For any $u = u_1 \cdots u_n$ is Σ^*, $u \models C_k^r\varphi$ iff the number of indices i with $u_i \cdots u_n \models \varphi$ is congruent to r modulo k. Similarly, let $u \models L_k^r$ iff $|u|$ is congruent to r modulo k. Let K and M be two subsets of the naturals. We denote by $\mathrm{FTL}(K, M)$ the extension of FTL by the modalities C_k^r, $0 \leq r < k$ and L_m^r, $0 \leq r < m$ with $k \in K$ and $m \in M$. Moreover, we denote by $\mathbf{FTL}(K, M)$ the class of all languages definable by the formulas in $\mathrm{FTL}(K, M)$.

For each $n \geq 1$, let Z_n denote a cyclic group of order n. Below we will denote by a a cyclic generator of Z_n and by 1 the identity element. The *division ideal* generated by a set M of naturals consists of all divisors of least common multiples of finite sets of naturals in M. When M is empty, the division ideal generated by M is $\{1\}$.

Theorem 6. *A language* L *belongs to* $\mathbf{FTL}(K, M)$ *iff* $Synt(L) = (S, A)$ *satisfies the following condition: There is some* m *in the division ideal generated by* M *such that every group contained in the submonoid of* S *generated by* A^m *is solvable whose order is a multiple of the prime divisors of the integers in* K.

Proof. We have $\mathbf{FTL}(K, M) = \mathbf{FTL}(\mathbf{K})$, where $\mathbf{K}$ consists of U^r, E and the mg-pairs $(Z_k, \{1, a\})$ and $(Z_m, \{a\})$, for all $k \in K$ and $m \in M$. Thus, by Theorem 3, $\mathbf{FTL}(\mathbf{K})$ consists of all languages whose syntactic mg-pair is in $\widehat{\mathbf{K}}$. It is shown in Ésik, Ito [6] that the syntactic mg-pair of a language belongs to this variety iff the condition described in the Theorem holds. □

See also Straubing, Therien, Thomas [12], Straubing [11].

For a class $\mathbf{K}$ of finite mg-pairs, let $\mathrm{FTL} + \mathbf{K}$ denote the extension of FTL by the modal operators corresponding to the languages recognizable by the mg-pairs in $\mathbf{K}$. Moreover, let $\mathbf{FTL} + \mathbf{K}$ denote the class of languages definable by the formulas in $\mathrm{FTL} + \mathbf{K}$.

Theorem 7. $\mathbf{FTL} + \mathbf{K}$ *is the class of all regular languages iff the following two conditions hold:*

1. *For each* m, $(Z_m, \{a\}) \in \widehat{\mathbf{K}}$.

2. *For each finite (nonabelian simple) group G there is an mg-pair (M, A) in* $\mathbf{K}$ *such that G divides M.*

Proof. First, by Theorem 5 and Theorem 2, it follows that a language belongs to $\mathbf{FTL} + \mathbf{K}$ iff its syntactic mg-pair is contained in the closed variety $\mathbf{V}$ generated by $\mathbf{K}$ and the aperiodics. Now, by (the dual of) a result proved in Dömösi, Ésik [4], $\mathbf{V}$ is the class of all finite mg-pairs iff the above two conditions hold. The result now follows by using Theorem 3, since a language is regular iff it is recognizable by a finite mg-pair. □

Corollary 8. *If* $\mathbf{FTL}+\mathbf{K}$ *is the class of all regular languages, then* $\mathbf{K}$ *is infinite.*

Example 3. Let $\mathbf{K}$ consist of the mg-pairs $(S_n, \{\pi, \rho\})$, $n \geq 3$, where S_n denotes the symmetric group of all permutations of the set $[n]$, and where π is a transposition and ρ is a cyclic permutation of $[n]$. Then $\mathbf{FTL} + \mathbf{K}$ is the class of all regular languages.

When $\mathbf{K}$ is a class of finite mg-pairs, let $\mathbf{FTL} + \mathbf{MOD} + \mathbf{K}$ denote the language class $\mathbf{FTL} + \mathbf{K}'$, where $\mathbf{K}' = \mathbf{K} \cup \{(Z_n, \{1, a\}) : n \geq 2\}$. The proof of the following result is similar to that of Theorem 7.

Theorem 8. *A language L belongs to* $\mathbf{FTL} + \mathbf{MOD} + \mathbf{K}$ *iff L is regular and for every finite nonabelian simple group G, if G divides the syntactic monoid of L, then G divides the monoid component of an mg-pair in* $\mathbf{K}$.

Given a formula φ of FTL over the alphabet Σ, let $\Diamond\varphi$ denote the formula $\mathtt{tt}\mathsf{U}\varphi$. Thus, for each word $u = u_1 \cdots u_n$ in Σ^*, $u \models \Diamond\varphi$ iff $u_1 \cdots u_n \models \varphi$, for some $i \in [n]$. In [3], Cohen, Perrin and Pin studied the expressive power of the restricted temporal logic RTL whose formulas over an alphabet Σ are constructed from the atomic formulas p_σ, $\sigma \in \Sigma$ by the X and $\Diamond$ modalities. Let $\mathbf{RTL}$ denote the class of languages definable by the formulas in RTL. Let U_1 denote a two-element semilattice (which may be identified with the mg-pair (U_1, U_1)).

Proposition 10. $\mathbf{RTL} = \mathbf{FTL}(\{U_1, E\})$.

The proof is based on the observation that U_1 is isomorphic to the syntactic monoid of the two-letter language $L = (a + b)^* b (a + b)^*$, and for any formula φ, $\Diamond\varphi$ is expressible as $L(a \mapsto \neg\varphi,\, b \mapsto \varphi)$.

Recall that a semigroup S is called $\mathbf{L}$-*trivial*, cf. Almeida [1], Eilenberg [5], if Green's $\mathbf{L}$-relation on S is the equality relation. Moreover, a semigroup S is *locally* $\mathbf{L}$-*trivial* iff for each idempotent e, the monoid eSe is $\mathbf{L}$-trivial. Accordingly, we call an mg-pair (M, A) locally $\mathbf{L}$-trivial if the subsemigroup of M generated by A is locally $\mathbf{L}$-trivial.

It follows from well-known facts (cf. [5, 1]) that a finite mg-pair is locally $\mathbf{L}$-trivial iff it belongs to the least closed variety containing U_1 and E_1, or U_1 and E. Thus, from Proposition 10 and Theorem 2 we may derive:

Theorem 9. Cohen, Perrin, Pin [3] *A language $L \subseteq \Sigma^*$ belongs to* **RTL** *iff L is regular and $Synt(L)$ is locally* **L***-trivial.*

Following the definitions of the language classes $\mathbf{FTL}(K, M)$, we may define the classes $\mathbf{RTL}(K, M)$. For lack of space we omit the proof of the following result.

Theorem 10. *A language L belongs to $\mathbf{RTL}(K, M)$ iff $Synt(L) = (S, A)$ satisfies the following condition: There is an integer m in the division ideal generated by M such that for each idempotent e of the subsemigroup T of S generated by A^m, it holds that eTe is a solvable group whose order is a multiple of the primes that divide the integers in K.*

References

1. J. Almeida, *Finite Semigroups and Universal Algebra*, World Scientific, 1994.
2. A. Baziramwabo, P. McKenzie and D. Therien, Modular temporal logic, in: *Proc. 1999 IEEE Conference LICS, Trento, Italy.*
3. J. Cohen, J.-E. Pin and D. Perrin, On the expressive power of temporal logic, *J. Computer and System Sciences*, 46(1993), 271–294.
4. P. Dömösi and Z. Ésik, Critical classes for the α_0-product, *Theoretical Computer Science*, 61(1988), 17–24.
5. S. Eilenberg, *Automata, Languages, and Machines*, vol. A and B, Academic Press, 1974 and 1976.
6. Z. Ésik and M. Ito, Temporal logic with cyclic counting and the degree of aperiodicity of finite automata, *Acta Cybernetica*, to appear.
7. Z. Ésik and K. G. Larsen, Regular languages definable by Lindström quantifiers, *BRICS Reports*, RS-02-20, April, 2002.
8. J. A. Kamp, Tense logic and the theory of linear order, Ph. D. Thesis, UCLA, 1968.
9. R. McNaughton and S. Papert, *Counter-Free Automata*, MIT Press, 1971.
10. A. Pnueli, The temporal logic of programs, in: *Proc. 18th FOCS, Providence, RI*, IEEE Press, 1977, 46–57.
11. H. Straubing, *Finite Automata, Formal Logic, and Circuit Complexity*, Birkhauser, 1994.
12. H. Straubing, D. Therien and W. Thomas, Regular languages defined with generalized quantifiers, *Information and Computation*, 118(1995), 289–301.
13. D. Therien and Th. Wilke, Temporal logic and semidirect products: An effective characterization of the until hierarchy, *DIMACS TR-96-28*, 1996.
14. P. Wolper, Temporal logic can be more expressive, *Information and Control*, 56(1983), 72–99.

Recent BRICS Report Series Publications

RS-02-47 Zoltán Ésik. ***Extended Temporal Logic on Finite Words and Wreath Product of Monoids with Distinguished Generators*****. December 2002. 16 pp. To appear in *6th International Conference, Developments in Language Theory*, DLT '02 Revised Papers, LNCS, 2002.**

RS-02-46 Zoltán Ésik and Hans Leiß. ***Greibach Normal Form in Algebraically Complete Semirings*****. December 2002. 43 pp. An extended abstract appears in Bradfield, editor, *European Association for Computer Science Logic: 16th International Workshop*, CSL '02 Proceedings, LNCS 2471, 2002, pages 135–150.**

RS-02-45 Jesper Makholm Byskov. ***Chromatic Number in Time*** $O(2.4023^n)$ ***Using Maximal Independent Sets*****. December 2002. 6 pp.**

RS-02-44 Zoltán Ésik and Zoltán L. Németh. ***Higher Dimensional Automata*****. November 2002. 32 pp. A preliminary version appears under the title *Automata on Series-Parallel Biposets* in Kuich, Rozenberg and Salomaa, editors, *5th International Conference, Developments in Language Theory*, DLT '01 Revised Papers, LNCS 2295, 2001, pages 217–227. This report supersedes the earlier BRICS report RS-01-24.**

RS-02-43 Mikkel Christiansen and Emmanuel Fleury. ***Using IDDs for Packet Filtering*****. October 2002. 25 pp.**

RS-02-42 Luca Aceto, Jens A. Hansen, Anna Ingólfsdóttir, Jacob Johnsen, and John Knudsen. ***Checking Consistency of Pedigree Information is NP-complete (Preliminary Report)*****. October 2002. 16 pp.**

RS-02-41 Stephen L. Bloom and Zoltán Ésik. ***Axiomatizing Omega and Omega-op Powers of Words*****. October 2002. 16 pp.**

RS-02-40 Luca Aceto, Willem Jan Fokkink, and Anna Ingólfsdóttir. ***A Note on an Expressiveness Hierarchy for Multi-exit Iteration*****. September 2002. 8 pp.**

RS-02-39 Stephen L. Bloom and Zoltán Ésik. ***Some Remarks on Regular Words*****. September 2002. 27 pp.**