FREE PRODUCTS OF INVERSE SEMIGROUPS II

by PETER R. JONES, STUART W. MARGOLIS†, JOHN MEAKIN† and JOSEPH B. STEPHEN

(Received 21 May, 1990; revised 24 December, 1990)

Let S and T be inverse semigroups. Their *free product* S inv T is their coproduct in the category of inverse semigroups, defined by the usual commutative diagram. Previous descriptions of free products have been based, like that for the free product of groups, on quotients of the free semigroup product S sgp T. In that framework, a set of canonical forms for S inv T consists of a transversal of the classes of the congruence associated with the quotient. The general result **[4]** of Jones and previous partial results **[3]**, **[5]**, **[6]** take this approach.

The approach here is to make use of the relationships among the presentations of S, T and S inv T. Recall that if X is a nonempty set and X^{-1} is a set of formal inverses of X then the *free inverse semigroup* FIS(X) over X is $(X \cup X^{-1})^+/\rho$, where $(X \cup X^{-1})^+$ is the free semigroup on $X \cup X^{-1}$ and ρ is the Vagner congruence (see **[10]**). If P is a relation on X then the *inverse semigroup with presentation* $(X \mid P)$, $\mathrm{Inv}\langle X \mid P\rangle$, is the quotient $(X \cup X^{-1})^+/\tau$, where τ is the congruence generated by $\rho \cup P$. (Clearly, $\mathrm{Inv}\langle X \mid P\rangle$ is also isomorphic to a quotient of the free inverse semigroup on X, by a suitable congruence.)

Now if the given inverse semigroups S and T are presented as $S = \mathrm{Inv}\langle X \mid P\rangle$ and $T = \mathrm{Inv}\langle Y \mid Q\rangle$, where X and Y are disjoint, then it is an exercise in universal algebra to verify that $S \text{ inv } T \cong \mathrm{Inv}\langle X \cup Y \mid P \cup Q\rangle$. The graph-theoretical techniques developed by Stephen to study presentations of inverse semigroups may then be used and this is the point of view that will be taken in this paper. These techniques originated with Munn's use of trees to study free inverse semigroups **[9]** and the current paper is therefore in a sense also a sequel to his.

Stephen's techniques are reviewed rather summarily in Section 1. For many further details the reader is referred to the paper **[12]** and the thesis **[13]** by Stephen (see also **[8]**). Section 2 treats in the abstract the particular types of graphs that appear in the construction of the Schützenberger automata for free products, the construction itself appearing in Section 3 (Theorem 3.4). Exactly which automata can be the Schützenberger automaton of an element of a free product is established in Section 4 (Theorem 4.1), thus providing a set of canonical forms for the product. From these forms, the canonical forms found by Jones in **[4]** may be quite easily found and proven unique (Theorem 4.5 and its corollary). Yet another set of canonical forms is also produced (Theorem 4.7 and its corollary). These are again graphical and are very similar to those previously given by Jones **[3]** for the free product of E-unitary inverse semigroups and by Margolis and Meakin **[7]** for the analogous product of inverse monoids, in the category of inverse monoids.

The paper is completed by a new application: the free product of two residually finite inverse semigroups is again residually finite.

We note that the methods of this paper may be extended in obvious ways to the free product of finitely many inverse semigroups, and by means of direct limits, to arbitrary

† Research supported by N.S.F. grant DMS 8702019.

families. In addition, some of the results may be modified so as to construct the Schützenberger automata for the free product of inverse monoids, within the category of inverse monoids. As remarked above, a special case was studied in [7], to which the reader is referred for a discussion of the differences between the two products.

For basic notions of semigroup theory see Clifford and Preston [1] and for inverse semigroups in particular, see Petrich [10]. For basic automata theory see Hopcroft and Ullman [2].

1. Inverse word graphs. Throughout this section, S is an inverse semigroup with fixed presentation $(X \mid P)$. A labelled digraph Γ over a nonempty set T, of labels, consists of a set $V(\Gamma)$, of vertices, and a set $E(\Gamma)$, of edges, where $E(\Gamma) \subseteq V(\Gamma) \times T \times V(\Gamma)$. An (*inverse*) *word graph* over X (more precisely, over $X \cup X^{-1}$) is a strongly connected digraph Γ whose edges are labelled from $X \cup X^{-1}$ in such a way that for any edge labelled by y there is an edge labelled by y^{-1} in the reverse direction. (In diagrams the edges labelled from X^{-1} are conventionally omitted.) A *path* in Γ is a nonempty sequence p of consecutive edges (whose vertices may be repeated). If the initial and terminal vertices of p are α and β, respectively, we write $p : \alpha \to \beta$. For $p = p_1 p_2 \ldots p_n$, the *word labelling* p is the string $y_1 y_2 \ldots y_n \in (X \cup X^{-1})^+$, where y_i labels p_i. The inverse path p^{-1} is then labelled by $y_n^{-1} \ldots y_2^{-1} y_1^{-1}$.

An *inverse automaton over* X is a triple $\mathscr{A} = (\alpha, \Gamma, \beta)$, where Γ is an inverse word graph and $\alpha, \beta \in V(\Gamma)$. (We may also term $\mathscr{A}$ a *birooted* word graph and if $\alpha = \beta$ we may call it *rooted*.) The *language* $L(\mathscr{A})$ of $\mathscr{A}$ consists of the set of words in $(X \cup X^{-1})^+$ that label paths in Γ from α to β.

An inverse automaton (and likewise a word graph) is *deterministic* if all edges directed from a vertex are labelled by different letters. A *homomorphism* of word graphs is a homomorphism of the underlying graph that preserves the labelling; a *homomorphism* of inverse automata is a homomorphism of the underlying word graphs that maps roots to roots. A V-*equivalence* on an inverse word graph Γ is an equivalence relation on $V(\Gamma)$. If η is such a relation then the V-*quotient* Γ/η is defined by $V(\Gamma/\eta) = V(\Gamma)/\eta$ and $E(\Gamma/\eta) = \{((v_1\eta), x, (v_2\eta)) \mid (v_1, x, v_2) \in E(\Gamma)\}$. V-equivalences on inverse automata are defined in the obvious way.

For $u \in (X \cup X^{-1})^+$, the Schützenberger graph $S\Gamma(u)$ of u (with respect to $(X \mid P)$) is the graph whose set of vertices is $R_{u\tau}$, the $\mathscr{R}$-class of the image of u in S; there is an edge labelled $y \in X \cup X^{-1}$ from $v\tau$ to $w\tau$ if $(vy)\tau = w\tau$. The *Schützenberger automaton* $\mathscr{A}(u)$ is the inverse automaton $((uu^{-1})\tau, S\Gamma(u), u\tau)$. By [12, Corollary 3.2],

$$L(\mathscr{A}(u)) = u\uparrow = \{v \in (X \cup X^{-1})^+ : v\tau \geq u\tau \text{ in } S\},$$

where $\geq$ is the natural partial order on S. Clearly, every Schützenberger graph and automaton is deterministic, a fact which will be used without comment. We may sometimes blur the terminology by referring to the Schützenberger graph or automaton of an element of S. This is justified by the following result from [12].

RESULT 1.1. *For* $S = \mathrm{Inv}\langle X \mid P\rangle = (X \cup X^{-1})^+/\tau$ *and* $u, v \in (X \cup X^{-1})^+$ *the following are equivalent*:

(a) $u\tau = v\tau$,
(b) $L(\mathscr{A}(u)) = L(\mathscr{A}(v))$,
(c) $u \in L(\mathscr{A}(v))$ *and* $v \in L(\mathscr{A}(u))$,
(d) $\mathscr{A}(u) \cong \mathscr{A}(v)$.

This result may be regarded as providing a set of canonical forms for S. Actual solution of the word problem for S depends on the decidability of these automata. We now present the procedure for constructing $\mathscr{A}(u)$ given by Stephen. This procedure is not in general effective but when it is it may be used to solve the word problem in S. For instance, the Schützenberger automata of free inverse semigroups are precisely the trees used by Munn [9] to solve the word problem. For various further applications, see [12].

We first extend the range of use of a definition in [12]. An inverse automaton $\mathscr{A}=(\alpha,\Gamma,\beta)$ over X is an *approximate automaton* for $u\in(X\cup X^{-1})^+$, relative to the given presentation, if (i) $u\in L(\mathscr{A})$ and (ii) $L(\mathscr{A})\subseteq u\uparrow$; thus if $\mathscr{A}$ is an approximate automaton for both u and v then $u\tau=v\tau$; the Schützenberger automaton of u is itself an approximate automaton for u. Call $\mathscr{A}$ an *approximate automaton* (relative to the presentation) if it is an approximate automaton for some $u\in(X\cup X^{-1})^+$. Call a word graph Γ an *approximate word graph* if for some $\alpha,\beta\in V(\Gamma)$, (α,Γ,β) is an approximate automaton. By [12, Theorem 3.1], this definition is independent of the choice of α and β.

RESULT 1.2 [12, Lemmas 2.3, 2.4, Theorem 2.5]. *Let $\mathscr{A}_1$ and $\mathscr{A}_2$ be inverse automata over X. If there exists a homomorphism from $\mathscr{A}_1$ to $\mathscr{A}_2$ then $L(\mathscr{A}_1)\subseteq L(\mathscr{A}_2)$. If $\mathscr{A}_1$ and $\mathscr{A}_2$ are deterministic and $L(\mathscr{A}_1)\subseteq L(\mathscr{A}_2)$ then there is a homomorphism from $\mathscr{A}_1$ to $\mathscr{A}_2$ and if $L(\mathscr{A}_1)=L(\mathscr{A}_2)$ then $\mathscr{A}_1\cong\mathscr{A}_2$.*

Let

$$\mathscr{A}_0\xrightarrow{\phi_1}\mathscr{A}_1\xrightarrow{\phi_2}\mathscr{A}_2\longrightarrow\dots\xrightarrow{\phi_n}\mathscr{A}_n\longrightarrow\dots \tag{1}$$

be a sequence of inverse automata, where the ϕ_i are homomorphisms. The direct limit of the sequence is defined in the usual way and is clearly again an inverse automaton over X.

LEMMA 1.3. *Let $w\in(X\cup X^{-1})^+$. The direct limit of a sequence* (1) *of approximate automata for w is again an approximate automaton for w.*

Proof. Let $\mathscr{A}$ be the direct limit of the sequence (1). It is easily verified that $L(\mathscr{A})=\bigcup\{L(\mathscr{A}_i):i\geq 1\}$, whence $w\in L(\mathscr{A})\subseteq w\uparrow$.

Two constructions are defined in [12] for an automaton $\mathscr{A}=(\alpha,\Gamma,\beta)$ over X: determinations and expansions. If the word graph Γ has two edges with common initial vertex and the same label, the automaton obtained from $\mathscr{A}$ by identifying their terminal vertices is an *elementary determination* of $\mathscr{A}$. The *determinized form* of $\mathscr{A}$ is the quotient of $\mathscr{A}$ by the least V-equivalence η on $\mathscr{A}$ such that $\mathscr{A}/\eta$ is deterministic. That a least such relation exists is proven in [13, Lemma 4.3]. (The determinized form of a word graph is defined similarly.) A *partial determination* is a quotient $\mathscr{A}/\eta_1$, where $\eta_1\subseteq\eta$. A similar definition appears in a paper by Stallings [11].

LEMMA 1.4. *Let $w\in(X\cup X^{-1})^+$. If $\mathscr{A}$ is an approximate automaton for w then so is any partial determination of $\mathscr{A}$.*

Proof. Let $\mathscr{B}$ be a partial determination of $\mathscr{A}$. Then $L(\mathscr{A})\subseteq L(\mathscr{B})$ by Result 1.2. By [12, Corollary 4.5], $L(\mathscr{B})\subseteq L(\mathscr{A})\uparrow\subseteq(u\uparrow)\uparrow=u\uparrow$, as required.

If $r=s$ is a relation in P and Γ has a path, from γ to δ, say, labelled by r but no such path labelled by s, "sew on" a new path from γ to δ, labelled by s. The resulting automaton is an *elementary expansion* of $\mathscr{A}$. An *expansion* of $\mathscr{A}$ is the automaton that results from simultaneously performing a set of elementary expansions (on $\mathscr{A}$ itself) and then determinizing. Denoting the undeterminized form by $E\mathscr{A}$, it is clear that $\mathscr{A}$ is embeddable in $E\mathscr{A}$ and that there is a homomorphism from $\mathscr{A}$ into the consequent expansion. Further, $L(E\mathscr{A})\tau \supseteq L(\mathscr{A})\tau$. If all possible elementary expansions are performed the result is the *full expansion* of $\mathscr{A}$.

Lemma 1.5. *Let $w \in (X \cup X^{-1})^+$. If $\mathscr{A}$ is an approximate automaton for w then so is any expansion of $\mathscr{A}$ and so is the undeterminized form $E\mathscr{A}$ of the expansion.*

Proof. Apply Lemma 1.4 and the remarks above.

An inverse automaton over X is *closed* if it is deterministic and no further expansions can be applied to it.

The procedure devised by Stephen is now presented. Let $w = u_1u_2 \ldots u_n$, where each $u_i \in (X \cup X^{-1})^+$ and let $\mathscr{A}_0(w)$ be the linear automaton of w:

$$\underset{\alpha}{\circ} \xrightarrow{u_1} \circ \xrightarrow{u_2} \circ \cdots\cdots\rightarrow \circ \xrightarrow{u_n} \underset{\beta}{\circ}$$

Let $\mathscr{A}_1(w)$ be the full expansion of $\mathscr{A}_0(w)$ and let $\phi_1 : \mathscr{A}_0(w) \rightarrow \mathscr{A}_1(w)$ be the homomorphism mentioned earlier. In general, let $\mathscr{A}_i(w)$ be the full expansion of $\mathscr{A}_{i-1}(w)$, with associated homomorphism $\phi_i : \mathscr{A}_{i-1}(w) \rightarrow \mathscr{A}_i(w)$.

The direct limit $\mathscr{D}(w)$ of the sequence

$$\mathscr{A}_0(w) \xrightarrow{\phi_1} \mathscr{A}_1(w) \xrightarrow{\phi_2} \mathscr{A}_2(w) \longrightarrow \ldots \xrightarrow{\phi_n} \mathscr{A}_n(w) \longrightarrow \ldots \tag{2}$$

is a deterministic inverse automaton over X which, by Lemma 1.3, is an approximate automaton for w and which is closed. The following result is a slight variation on [**12**, Theorem 5.10].

Result 1.6. *Let $w \in (X \cup X^{-1})^+$. If $\mathscr{A}$ is a closed approximate automaton for w then $\mathscr{A}$ is isomorphic to the Schützenberger automaton for w.*

The main tool for application in the sequel is now evident.

Theorem 1.7. *Let $w \in (X \cup X^{-1})^+$. The Schützenberger automaton $\mathscr{A}(w)$ is isomorphic, as an inverse automaton over X, to the direct limit of the sequence* (2) *of iterated full expansions of the linear graph of w.*

(If the presentation $(X \mid P)$ is finite, the sequence (2) will consist of finite automata. If the sequence should stabilize, this procedure thus provides an effective construction for the Schützenberger graph of w. If all such sequences stabilize, the word problem for the inverse semigroup thus presented is therefore solved by Result 1.1. In [**12**] a "confluence lemma" is proved, implying that any closed automaton that is obtained from the linear graph of w by a sequence of expansions and determinations is isomorphic to the Schützenberger graph of w. A more general result yet is contained in [**13**]. However, we shall not need any of these further observations.)

This introduction is completed with a method for forming a product of inverse automata $(\alpha_1, \Gamma_1, \beta_1)$ and $(\alpha_2, \Gamma_2, \beta_2)$ over X. This product $(\alpha_1, \Gamma_1, \beta_1) \times (\alpha_2, \Gamma_2, \beta_2)$ is the inverse automaton $(\alpha_0, \Gamma_0, \beta_0)$, whose underlying word graph is $\Gamma_0 = (\Gamma_1 \cup \Gamma_2)/\theta$, where θ is the least V-equivalence on $\Gamma_1 \cup \Gamma_2$ that identifies the vertices β_1 and α_2, and α_0

and β_0 are the images of α_1 and β_2, respectively, in that quotient graph. Clearly, the product of two deterministic automata $(\alpha_1, \Gamma_1, \beta_1)$ and $(\alpha_2, \Gamma_2, \beta_2)$ need not again be deterministic. However, if Γ_1 and Γ_2 are subgraphs of some deterministic word graph Γ, and if $\beta_1 = \alpha_2$ in Γ then Γ_0 is again a subgraph of Γ and the product is therefore once again deterministic.

RESULT 1.8 [**12**, Lemma 5.2]. *Let $\mathscr{A}_1$ and $\mathscr{A}_2$ be inverse automata over X. If $\mathscr{A}_1$ and $\mathscr{A}_2$ are approximate automata for u_1 and u_2, respectively, in $(X \cup X^{-1})^+$ then $\mathscr{A}_1 \times \mathscr{A}_2$ is an approximate automaton for u_1u_2.*

2. Word graphs over $X \cup Y$**.** Let $(X \mid P)$ and $(Y \mid Q)$ be disjoint inverse presentations. In this section we study the abstract properties of the word graphs that appear in the sequel. The edges of a word graph Γ over $X \cup Y$ may be considered "colored" by one of two "colors"—one for $X \cup X^{-1}$ and one for $Y \cup Y^{-1}$. A subgraph of Γ is (*edge*-) *monochromatic* if all its edges have the same color. (In general, then, word graphs over $X \cup Y$ are "dichromatic".) A path is *monochromatic* if its underlying subgraph is. A *lobe* of Γ is a maximal monochromatic connected subgraph of Γ. Clearly the lobes of Γ partition its edge set $E(\Gamma)$ and each lobe is a word graph over either X or Y. A vertex of Γ that belongs to more than one lobe (clearly, exactly two) is called an *intersection* (or dichromatic) vertex.

A path in a digraph is *simple* if it contains no repeated vertex, other than perhaps its first and last, in which case it is a *simple cycle.*

A word graph Γ over $X \cup Y$ is called *cactoid* if

(i) Γ has finitely many lobes, and

(ii) every simple cycle of Γ is monochromatic.

We shall prove in Section 3 that the Schützenberger graphs of the elements of $\mathrm{Inv}\langle X \cup Y \mid P \cup Q\rangle$ are, up to isomorphism, the cactoid graphs over $X \cup Y$ each of whose lobes is the Schützenberger graph of an element of $\mathrm{Inv}\langle X \mid P\rangle$ or of $\mathrm{Inv}\langle Y \mid Q\rangle$. This section is devoted to the study of cactoid graphs. Throughout the remainder of the section Γ will denote such a graph.

Let $p : \gamma \to \delta$ be a path in Γ. A *switchpoint* of p is a vertex that is common to successive edges of different color. (Thus each switchpoint of p is an intersection point of Γ, but p may pass through an intersection point without changing color.) The *switchpoint sequence* of p is then the (possibly empty) sequence of switchpoints which p traverses, in order. The path p may be factored as a product of monochromatic subpaths. The sequence of colors of those subpaths is its *color sequence.*

LEMMA 2.1. *Let γ and δ be distinct vertices of Γ. Any two simple paths from γ to δ traverse the same switchpoint sequence and the same color sequence.*

Proof. The proof is by induction on the number of common vertices of the paths. Suppose $p, q : \gamma \to \delta$ are simple paths with no common vertices other than γ and δ. Then the composite path pq^{-1} is a simple cycle in Γ whence, by hypothesis, monochromatic. Thus p and q have the same color and the same (empty) switchpoint sequences.

On the other hand, if p and q have a common vertex π, other than γ and δ, then p and q may be factored, as p_1p_2 and q_1q_2 respectively, where $p_1, q_1 : \gamma \to \pi$ and $p_2, q_2 : \pi \to \delta$. Now the induction hypothesis gives the requisite result, for p_1 and q_1 end in the same color and p_2 and q_2 begin in the same color; so π itself can be a switchpoint for p if and only if it is for q.

(In fact the statement in the lemma is equivalent to the monochromaticity of simple cycles in Γ.)

COROLLARY 2.2. *Distinct lobes of* Γ *possess at most one common vertex.*

Proof. Suppose lobes Λ_1 and Λ_2 possess common vertices γ, δ, $\gamma \neq \delta$. There are simple paths p and q, in Λ_1 and Λ_2, respectively, from γ to δ. By the preceding lemma, p and q have the same color sequence and therefore the same color. Thus $\Lambda_1 = \Lambda_2$.

For distinct vertices γ, δ of Γ, the *switchpoint sequence from* γ *to* δ is the switchpoint sequence of some (any) simple path from γ to δ.

An automaton will be termed *cactoid* if its underlying word graph is. Let $\mathcal{A} = (\alpha, \Gamma, \alpha)$ be a cactoid rooted automaton. Let γ be a vertex of Γ distinct from α. The *switchpoint sequence of* γ is that from α to γ. Let l be the length of the switchpoint sequence of γ. The *norm* of γ (with respect to the root α) is defined by:

$$\|\gamma\| = \begin{cases} 0, & \text{if } \gamma = \alpha, \\ 1 + l, & \text{otherwise.} \end{cases}$$

The *nodes* of $\mathcal{A}$ are the intersection vertices of Γ, together with the root α.

LEMMA 2.3. *Each lobe* Λ *of* Γ *possesses a unique vertex* λ_Λ *of least norm and* λ_Λ *is a node of* $\mathcal{A}$. *For every other vertex* γ *of* Λ, $\|\gamma\| = \|\lambda_\Lambda\| + 1$.

Proof. Suppose $\alpha \in \Lambda$. Then $\|\alpha\| = 0$ and, for every other vertex γ of Λ, $\|\gamma\| = 1$. Otherwise, let λ be a vertex of least norm in Λ. Let p be a simple path from α to λ. Then the final edge in p does not belong to Λ, for the final switchpoint μ, say, of p would then belong to Λ but have norm $\|\lambda\| - 1$. Hence λ is a node of $\mathcal{A}$. Let γ be any other vertex of Λ and let q be a simple path from λ to γ in Λ. It follows from Lemma 2.1 that the composite path pq is a simple path from α to γ that has λ as its final switchpoint. Thus $\|\gamma\| = \|\lambda\| + 1$.

The node λ_Λ defined in the lemma will be called the *root* of Λ.

COROLLARY 2.4 (to the proof). *The root* α *of* $\mathcal{A}$ *is the root of two distinct lobes if it is an intersection vertex, and is the root of a unique lobe otherwise. Every other node is the root of a unique lobe.*

Proof. The first sentence is clear from $\|\alpha\| = 0$. Let γ be any other node and let p be a simple path from α to γ. Then γ is the root of the lobe to which the last edge of p does not belong and is not the root of the lobe to which it does belong.

3. The construction. Let $(X \mid P)$ and $(Y \mid Q)$ be disjoint presentations and let $S = \mathrm{Inv}\langle X \mid P\rangle$ and $T = \mathrm{Inv}\langle Y \mid Q\rangle$. Let $w \in (X \cup X^{-1} \cup Y \cup Y^{-1})^+$. We describe a procedure for obtaining the Schützenberger automaton of w (relative to the presentations of S and of T) from the linear automaton of w in a finite number of steps. The typical step simply replaces a single lobe by the Schützenberger graph of the element of $S \cup T$ that it represents. Thus effective computation of $\mathcal{A}(w)$ relies only on effective computation of the Schützenberger graphs of the elements of S and of T.

Denote the linear automaton $\mathscr{A}_0(w)$ of w (see Section 1) by $(\alpha_w, \Gamma_w, \beta_w)$. The graph Γ_w is clearly cactoid. Let the nodes of $(\alpha_w, \Gamma_w, \alpha_w)$ be $\alpha_w = \lambda_0, \lambda_1, \ldots, \lambda_n$, say, $n \geq 0$. Thus for each i, $\|\lambda_i\| = i$ and λ_i is the root of a unique lobe Λ_i. Put $\lambda_{n+1} = \beta_w$. With each lobe Λ_i there is associated the inverse automaton $(\lambda_i, \Lambda_i, \lambda_{i+1})$ which is the linear automaton of some word $w^{(i)}$ in either $(X \cup X^{-1})^+$ or $(Y \cup Y^{-1})^+$ and is therefore an approximate automaton for $w^{(i)}$ over X or Y. It follows that each rooted lobe $(\lambda_i, \Lambda_i, \lambda_i)$ is also an approximate automaton (for $w^{(i)}(w^{(i)})^{-1}$).

We now define a construction that, when applied to the linear automaton of w, will yield its Schützenberger automaton in finitely many steps. Let $\mathscr{A} = (\alpha, \Gamma, \beta)$ by any cactoid inverse automaton over $X \cup Y$ with the property that each lobe is an approximate graph (over X or over Y).

Suppose $\mathscr{A}$ is not closed. Then either there are two edges of Γ with the same label and same initial vertex, or there is a path in Γ labelled by one side of some relation in $P \cup Q$ but no path between the same vertices labelled by the other side. In the former case, both edges are contained in the same lobe of Γ. In the latter case, since $P \subseteq (X \cup X^{-1})^+ \times (X \cup X^{-1})^+$ and $Q \subseteq (Y \cup Y^{-1})^+ \times (Y \cup Y^{-1})^+$, this path is again contained in some lobe of Γ. In either case, denote the lobe by Λ, with root $\lambda = \lambda_\Lambda$. Without loss of generality we may suppose Λ is colored from X, whence $(\lambda, \Lambda, \lambda)$ is an approximate automaton for some $u \in (X \cup X^{-1})^+$. However, by assumption, Λ is not closed with respect to $(X \mid P)$. By Theorem 1.7, there is a homomorphism ϕ from $(\lambda, \Lambda, \lambda)$ to the Schützenberger automaton $\mathscr{A}(u)$ of u with respect to $(X \mid P)$. Thus $\mathscr{A}(u) = (\lambda^*, \Lambda^*, \lambda^*)$, where $\Lambda^* = S\Gamma(u)$, the Schützenberger graph of u with respect to $(X \mid P)$, and $\lambda^* = \lambda\phi$.

Construct a new automaton $\mathscr{A}^*$ from $\mathscr{A}$ by replacing Λ by Λ^*, the Schützenberger graph of u over X. Formally, let $\Gamma^* = (\Gamma \cup \Lambda^*)/\kappa$, where κ is the least V-equivalence on $\Gamma \cup \Lambda^*$ that identifies each vertex of Λ with its image in Λ^*. (Another viewpoint on this construction is presented below.)

The homomorphism ϕ from Λ to Λ^* extends to a homomorphism ϕ^* of Γ to Γ^*; put $\alpha^* = \alpha\phi^*$, $\beta^* = \beta\phi^*$ and $\mathscr{A}^* = (\alpha^*, \Gamma^*, \beta^*)$. Clearly κ identifies only vertices within Λ (including, possibly, intersection vertices with other lobes). Thus ϕ^* is injective on lobes other than Λ; we will identify these remaining lobes with their images in Γ^*. The lobes of Γ^* therefore comprise: $\Lambda^* = S\Gamma(u)$; those lobes of Γ that did not intersect Λ; and lobes that were formed by the amalgamation of lobes of Γ by identification, under κ, of their intersection points with Λ.

Whilst "replacing the lobe by its Schützenberger graph" is the way this construction is intended to be viewed, an alternative viewpoint will simplify its validation. Let $\mathscr{L}_0$ denote the automaton $(\lambda, \Lambda, \lambda)$. By hypothesis, $\mathscr{L}_0$ is an approximate automaton for u. Hence, by Theorem 1.7, $\mathscr{A}(u)$ is the direct limit of the sequence

$$\mathscr{L}_0 \xrightarrow{\phi_1} \mathscr{L}_1 \xrightarrow{\phi_2} \mathscr{L}_2 \longrightarrow \ldots \xrightarrow{\phi_n} \mathscr{L}_n \longrightarrow \ldots \tag{3}$$

of iterated full expansions of $\mathscr{L}_0$ with respect to $(X \mid P)$. For each i, let $\mathscr{L}_i = (\lambda_i, \Lambda_i, \lambda_i)$. Then $\mathscr{A}^*$ is the direct limit of the sequence

$$\mathscr{A}_0 \xrightarrow{\phi_1^*} \mathscr{A}_1 \xrightarrow{\phi_2^*} \mathscr{A}_2 \longrightarrow \ldots \xrightarrow{\phi_n^*} \mathscr{A}_n \longrightarrow \ldots, \tag{4}$$

https://doi.org/10.1017/S0017089500008442 Published online by Cambridge University Press

where $\mathscr{A}_0 = \mathscr{A}$ and $\mathscr{A}_i$ is obtained from $\mathscr{A}_{i-1}$ by replacing the lobe Λ_{i-1} by Λ_i in $\mathscr{A}_{i-1}$, and ϕ_i^* is the induced homomorphism from $\mathscr{A}_{i-1}$ to $\mathscr{A}_i$.

Put $\mathscr{A}_i = (\alpha_i, \Gamma_i, \beta_i)$. The replacement of Λ_{i-1} by Λ_i involves an expansion (see Section 1) of $\mathscr{A}_{i-1}$: first all possible elementary expansions within Λ_{i-1} are performed to obtain the intermediate graph $E\Gamma_{i-1}$ and the intermediate automaton $E\mathscr{A}_{i-1}$. Then the partial determination of $E\mathscr{A}_{i-1}$ that determinizes $E\Lambda_{i-1}$ (to obtain Λ_i) is performed.

Proposition 3.1. *If $\mathscr{A}$ is an approximate automaton for w then so is $\mathscr{A}^*$.*

Proof. This follows from Lemmas 1.3 and 1.5.

Proposition 3.2. *The automaton $\mathscr{A}^*$ is cactoid.*

Proof. We use the original definition of $\mathscr{A}^*$. First observe that Γ^* has at most as many lobes as Γ has. Now let p be a simple cycle in its word graph Γ^*. Suppose p contains no vertex of the lobe Λ^*. Then (using the identification of lobes of $\mathscr{A}$ disjoint from Λ with their images in Γ^*), p is a simple path in Γ and is therefore monochromatic.

Next, suppose p contains a vertex of Λ^* but is not a path in that lobe. Without loss of generality, p contains nodes γ and δ of Λ^* (possibly equal) such that the subpath p_1 from γ to δ contains no vertex of Λ^* other than its endpoints. This subpath is then the image of a simple path q_1 from a preimage γ_1 of γ to a preimage δ_1 of δ in Γ. If γ_1 and δ_1 are distinct then there is a simple path within Λ from the former to the latter; its composition with q_1 forms a simple dichromatic cycle in Γ, contradicting the hypothesis that Γ is cactoid. Thus $\gamma_1 = \delta_1$ and q_1 is a simple cycle in Γ, whence monochromatic. Hence p_1 is monochromatic. Moreover, since $\gamma = \delta$ and p is simple, $p = p_1$. The only remaining possibility is that p lies within Λ^* and is therefore again monochromatic.

Proposition 3.3. *If each lobe of Γ is an approximate graph (over S or over T) then so is each lobe of Γ^*.*

Proof. The result is true of the new lobe Λ^*, for it is the Schützenberger graph of u. Any remaining lobe that was not just a lobe of Γ was formed from lobes $M_1, \ldots, M_k$ of Γ of the same color by amalgamating over a common vertex, μ say, belonging to Λ. By hypothesis, for each j, (μ, M_j, μ) is an approximate automaton. The amalgamated lobe M is then isomorphic to the underlying word graph of $(\mu, M_1, \mu) \times \ldots \times (\mu, M_k, \mu)$ which is an approximate automaton by Result 1.8.

The construction $\mathscr{A} \to \mathscr{A}^*$ may be iterated, beginning with the linear automaton $\mathscr{A}_0(w)$. The three preceding propositions verify that the hypotheses are valid at each iteration. This yields a sequence

$$\mathscr{A}_0(w), \mathscr{A}_0^*(w), \mathscr{A}_0^{**}(w) \ldots$$

At each stage the number of lobes either remains the same (if no identification of nodes occurs) or decreases. After finitely many iterations, therefore, no further identifications of nodes may occur. Thus each further iteration results in the closure of a single lobe, leaving the remaining lobes untouched. Since the automata have finitely many lobes, the process terminates in a closed automaton which is approximate, relative to $(X \cup Y \mid P \cup Q)$. An application of Lemma 1.6 completes the proof of our main theorem.

THEOREM 3.4. *Let $(X \mid P)$ and $(Y \mid Q)$ be disjoint presentations of inverse semigroups S and T, respectively. Let $w \in (X \cup X^{-1} \cup Y \cup Y^{-1})^+$. Beginning with the linear automaton of w, iterated application of the above construction, at each step replacing a single non-closed lobe with the associated Schützenberger graph, relative to S or to T, terminates in finitely many steps in the Schützenberger automaton of w.*

COROLLARY 3.5. *If there is an effective procedure for constructing the Schützenberger automata of elements of S and of T, there is such a procedure for constructing the automata for the elements of S* inv *T.*

4. Canonical forms. The main result of this section is the following theorem.

THEOREM 4.1. *Let $(X \mid P)$ and $(Y \mid Q)$ be disjoint presentations of inverse semigroups S and T, respectively. The Schützenberger automata of the elements of S inv $T = \mathrm{Inv}\langle X \cup Y \mid P \cup Q\rangle$ are precisely (a transversal of) the cactoid inverse automata over $X \cup Y$ each of whose lobes is a Schützenberger graph over either X or Y.*

This theorem provides a set of canonical forms for S inv T. Its proof leads naturally to a derivation of the canonical form of Jones (Theorem 4.5) in terms of words in the free semigroup product S sgp T and to an alternative canonical form by trees labelled from S and T.

That every Schützenberger automaton has the stated properties is immediate from Propositions 3.2 and 3.3 and the remarks prior to Theorem 3.4. Conversely, let $\mathcal{A} = (\alpha, \Gamma, \beta)$ be a cactoid automaton over $X \cup Y$ each of whose lobes is a Schützenberger graph over either X or Y. Let Λ be a lobe of $\mathcal{A}$ with root λ distinct from α. Define the *extended switchpoint sequence* of λ to be the uniquely defined sequence $\alpha = \lambda_0, \lambda_1, \ldots, \lambda_n = \lambda$, where the subsequence $\lambda_1, \ldots, \lambda_{n-1}$ is the switchpoint sequence of λ. Let $\Lambda^{(0)}, \ldots, \Lambda^{(n-1)}$ be the sequence of distinct lobes containing the edges of some simple path from α to λ; let $\Lambda^{(n)} = \Lambda$.

The *branch to* Λ, $\mathrm{br}(\Lambda)$, is the union of the subgraphs $\Lambda^{(0)}, \ldots, \Lambda^{(n)}$ and its *branch automaton* is $(\alpha, \mathrm{br}(\Lambda), \lambda)$. Its *rooted branch automaton* is $(\alpha, \mathrm{br}(\Lambda), \alpha)$. If $\lambda = \alpha$, put $\mathrm{br}(\Lambda) = \Lambda$ and define the branch automata similarly.

From the definition of product of automata (end of Section 1), it is clear that

$$(\alpha, \mathrm{br}(\Lambda), \lambda) \cong (\alpha, \Lambda^{(0)}, \lambda_1) \times \ldots \times (\lambda_{n-1}, \Lambda^{(n-1)}, \lambda) \times (\lambda, \Lambda, \lambda)$$

and that

$$(\alpha, \mathrm{br}(\Lambda), \alpha) \cong (\alpha, \mathrm{br}(\Lambda), \lambda) \times (\lambda, \mathrm{br}(\Lambda), \alpha).$$

By hypothesis, for $i = 0, \ldots, n-1$, there exists $u_i \in (X \cup X^{-1})^+ \cup (Y \cup Y^{-1})^+$ such that $\mathcal{A}(u_i) = (\lambda_i, \Lambda^{(i)}, \lambda_{i+1})$; and there exists u_n such that $\mathcal{A}(u_n) = (\lambda, \Lambda, \lambda)$. Put $u(\Lambda) = u_0 \ldots u_n$.

PROPOSITION 4.2. *The branch automaton $(\alpha, \mathrm{br}(\Lambda), \lambda)$ is the Schützenberger automaton of the product $u(\Lambda)$; the rooted branch automaton $(\alpha, \mathrm{br}(\Lambda), \alpha)$ is the Schützenberger automaton of the product $u(\Lambda)u(\Lambda)^{-1}$.*

Proof. By Result 1.8, $(\alpha, \mathrm{br}(\Lambda), \lambda)$ is an approximate automaton for $u(\Lambda)$; but, since Γ is closed, so is $\mathrm{br}(\Lambda)$ and the first statement follows from Result 1.6. The second statement is similar.

PROPOSITION 4.3. *The rooted automaton* (α, Γ, α) *is isomorphic with the product, in any order, of the rooted branch automata* $(\alpha, \mathrm{br}(\Lambda), \alpha)$ *over the set of lobes of* Γ. *Hence,* (α, Γ, α) *is the Schützenberger automaton of* $\prod \{u(\Lambda)u(\Lambda)^{-1} : \Lambda \text{ a lobe of } \Gamma\}$, *the product taken in any order.*

Proof. The first statement follows from a straightforward induction argument (note the comments preceding Result 1.8). The second uses the method of the preceding proposition.

Finally we take account of the vertex β. If $\beta = \alpha$, define $r(\beta) = 1$. Otherwise, let $\alpha = \beta_0, \ldots, \beta_m = \beta$ be its extended switchpoint sequence. Let $M^{(0)}, \ldots, M^{(m-1)}$ be the sequence of distinct lobes containing the edges of some simple path from α to β. For $i = 0, \ldots, m-1$, let r_i be a word in $(X \cup X^{-1})^+ \cup (Y \cup Y^{-1})^+$ whose Schützenberger automaton is $(\beta_i, M^{(i)}, \beta_{i+1})$. Put $r(\beta) = r_0 \ldots r_{m-1}$ (cf. the definition of $u(\Lambda)$ above).

PROPOSITION 4.4. *For* $\beta \neq \alpha$, *the automaton* $\mathcal{A} = (\alpha, \Gamma, \beta)$ *is isomorphic with the product*

$$(\alpha, \Gamma, \alpha) \times (\beta_0, M^{(0)}, \beta_1) \times \ldots \times (\beta_{m-1}, M^{(m-1)}, \beta).$$

Hence, in every case, $\mathcal{A}$ *is the Schützenberger automaton of the word*

$$\prod \{u(\Lambda)u(\Lambda)^{-1} : \Lambda \text{ a lobe of } \Gamma\} r(\beta) \in (X \cup X^{-1} \cup Y \cup Y^{-1})^+.$$

Proof. The isomorphism with the specified product is clear. The second statement follows from Proposition 4.3, by the same techniques.

This completes the proof of Theorem 4.1, characterizing the Schützenberger automata of the element of $S \operatorname{inv} T$ and therefore, in conjunction with Result 1.1, providing a set of canonical forms for the free product. Of course this form is dependent on the particular presentations provided for S and T. An alternative "presentation-independent" graphical canonical form will be presented later in this section. Next, however, we show that Jones' canonical forms [**4**] can be easily derived from Theorem 4.1, using the propositions in its proof.

We briefly review some notation and terminology from [**4**]. Elements of the free semigroup product $S \operatorname{sgp} T$ may be represented uniquely as words of the form $v_1 \ldots v_n$, where the letters belong alternately to S and to T. Since $S \operatorname{inv} T$ is generated, as an inverse semigroup, by S and T, we may view $S \operatorname{inv} T$ as a subset of $S \operatorname{sgp} T$, when convenient. A word in $S \operatorname{sgp} T$ is *left reduced* if its last letter, but no other, is an idempotent. A *canonical set* for $S \operatorname{inv} T$ is a finite nonempty set A of left reduced words such that

(i) A is prefix-closed, i.e. if $v_1 \ldots v_n \in A$ then $v_1 \ldots (v_i v_i^{-1}) \in A$ for $i = 1, \ldots, n$, and

(ii) A has unique last letters, i.e. if $v_1 \ldots v_n$, $w_1 \ldots w_n \in A$ and v_1, w_1 belong to the same factor S or T then $v_i = w_i$ for all $i < n$ implies $v_n = w_n$.

An *associate* of such a set A is either the empty word 1 or a reduced word $a = a_1 \ldots a_n$ (with no a_i idempotent) such that $a_1 \ldots (a_n a_n^{-1}) \in A$. A *canonical pair* (A, a) for $S \operatorname{inv} T$ consists of a canonical set A and an associate a of A.

If $a = a_1 \ldots a_n \in S \operatorname{sgp} T$, put $aa^{-1} = a_1 \ldots (a_n a_n^{-1}) \ldots a_1^{-1}$; if A is a finite nonempty subset of $S \operatorname{sgp} T$, put $\varepsilon(A) = \prod \{aa^{-1} : a \in A\}$, where the product is taken is some fixed order. (Considered as an element of $S \operatorname{inv} T$, the order in which the product is taken is immaterial.)

One procedure for associating a canonical pair with an element of $S \operatorname{inv} T$ is presented in **[4]**. Our procedure here is based on Theorem 4.1 and its proof. Thus we must first choose presentations $(X \mid P)$ and $(Y \mid Q)$ for S and T respectively. (However the end result will be independent of these choices.) Let $\mathscr{A} = (\alpha, \Gamma, \beta)$ be a Schützenberger automaton for $S \operatorname{inv} T$, according to Theorem 4.1. For each lobe Λ, with root λ, define $u(\Lambda) = u_1 \ldots u_n$ as above (before Proposition 4.2) and let $\bar{u}(\Lambda) = \overline{u_1} \ldots \overline{u_n}$, where $\overline{u_i}$ is the image of u_i in S or in T, depending on whether u_i is in $(X \cup X^{-1})^+$ or in $(Y \cup Y^{-1})^+$. By Proposition 4.2, $(\alpha, \operatorname{br}(\Lambda), \lambda) \cong \mathscr{A}(\overline{u_1} \ldots \overline{u_n})$.

From the definition of $u(\Lambda)$, it is clear that $\bar{u}(\Lambda)$ is a left reduced word in $S \operatorname{sgp} T$ and that for $i \le n$, $\overline{u_1} \ldots (\overline{u_i}\overline{u_i}^{-1}) = \bar{u}(\Lambda^{(i)})$. Hence the set $A(\mathscr{A}) = \{\bar{u}(\Lambda) : \Lambda \text{ a lobe of } \Gamma\}$ is prefix-closed. Let Λ, M be lobes of Γ, with roots λ, μ, respectively, and let $u(\Lambda) = u_1 \ldots u_n$, $u(M) = v_1 \ldots v_m$. Suppose $m = n$, $\overline{u_1}$, $\overline{v_1}$ both belong to the same factor S or T and $\overline{u_i} = \overline{v_i}$, $i < n$. Then a simple induction establishes that u_i and v_i represent the same lobes, for $i < n$, whence Λ and M have the same intersection point with the lobe represented by u_{n-1}, that is, $\lambda = \mu$. Therefore $\mathscr{A}(u_n) = (\lambda, \Lambda, \lambda) = \mathscr{A}(v_n)$ and $\overline{u_n} = \overline{v_n}$. Hence $A(\mathscr{A})$ has unique last letters.

Finally, recall the definition of $r(\beta)$ (before Proposition 4.4), define $\overline{r(\beta)}$ as for $\bar{u}(\Lambda)$ and denote it by $r(\mathscr{A})$. It is easily verified that $r(\mathscr{A})$ is a reduced word in $S \operatorname{sgp} T$ and is an associate of $A(\mathscr{A})$.

THEOREM 4.5. *The map $\mathscr{A} \to (A(\mathscr{A}), r(\mathscr{A}))$ is a 1–1 correspondence from the set of isomorphism classes of Schützenberger automata for $S \operatorname{inv} T$, with respect to $(X \cup Y \mid P \cup Q)$, to the set of canonical pairs for $S \operatorname{inv} T$, with inverse $(A, r) \to \mathscr{A}(\varepsilon(A)r)$.*

Before completing the proof of the theorem we draw its main conclusion.

COROLLARY 4.6. *The elements of $S \operatorname{inv} T$ are uniquely representable in the form $\varepsilon(A)r$, where (A, r) is a canonical pair for $S \operatorname{inv} T$.*

Proof. Choose presentations as above. Any element of $S \operatorname{inv} T$ has the form $w\tau$ for some $w \in (X \cup X^{-1} \cup Y \cup Y^{-1})^+$. Put $A = A(\mathscr{A}(w))$ and $r = r(\mathscr{A}(w))$. Then, by the theorem, (A, r) is a canonical pair and, regarding $\varepsilon(A)r$ as a member of $S \operatorname{inv} T$, we have $\mathscr{A}(w\tau) \cong \mathscr{A}(\varepsilon(A)r)$, so that $w\tau = \varepsilon(A)r$ in $S \operatorname{inv} T$. Uniqueness follows from the theorem and Theorem 4.1.

That the representation $\varepsilon(A)r$ is independent of the choice of presentations for S and T is immediate from the uniqueness of the representation.

Proof of Theorem 4.5. It has already been demonstrated that if $\mathscr{A}$ is a Schützenberger automaton for $S \operatorname{inv} T$ then $(A(\mathscr{A}), r(\mathscr{A}))$ is a canonical pair. Further, by Proposition 4.4, $\mathscr{A} \cong \mathscr{A}(\varepsilon(A(\mathscr{A}))r(A(\mathscr{A})))$. It remains to prove that the map $(A, r) \to \mathscr{A}(\varepsilon(A)r)$ is indeed the requisite inverse. In so doing we actually construct $\mathscr{A}(\varepsilon(A)r)$.

First let $v = v_1 \ldots v_n$ be a left reduced word in $S \operatorname{sgp} T$ and $\mathscr{A} = \mathscr{A}(v_1) \times \ldots \times \mathscr{A}(v_n)$. It is clear that since for each $i < n$, v_i is a nonidempotent and the successive terms alternate between S and T, the lobes of $\mathscr{A}$ are just (isomorphic copies of)

$\mathscr{A}(v_1), \ldots, \mathscr{A}(v_n)$ and, in fact, $\mathscr{A} \cong (\alpha, \mathrm{br}(\mathscr{A}(v_n)), \lambda)$, where α is the root of $\mathscr{A}(v_1)$ and λ is the root of $\mathscr{A}(v_n)$. By Proposition 4.1 and the remarks that precede it, $\mathscr{A} = \mathscr{A}(v_1 \ldots v_n)$ (interpreting the product in $S \operatorname{inv} T$), and $(\alpha, \mathrm{br}(\mathscr{A}(v_n)), \alpha) \cong \mathscr{A}(v_1 \ldots (v_n v_n^{-1}) \ldots v_1^{-1}) = \mathscr{A}(vv^{-1})$.

Now let A be a canonical set for $S \operatorname{inv} T$. Form the product of the rooted automata $\mathscr{A}(vv^{-1})$, $v \in A$, as just constructed. The result is a rooted automaton (α, Γ, α), say, whose branches are clearly (isomorphic copies of) the individual automata in the product. By Proposition 4.3, $(\alpha, \Gamma, \alpha) \cong \mathscr{A}(\varepsilon(A))$, where the product $\varepsilon(A)$ is interpreted in $S \operatorname{inv} T$.

Finally, if $r = 1$, let $\beta = \alpha$. Otherwise, $r = r_1 \ldots r_m$, say, where $r_1 \ldots (r_m r_m^{-1}) \in A$. Then for some lobe Λ of Γ, with root λ, $(\lambda, \Lambda, \lambda) = \mathscr{A}(r_m r_m^{-1})$, (with respect to $(X \mid P)$ or $(Y \mid Q)$). Since $r_m \mathscr{R} r_m r_m^{-1}$, we may let β be the vertex of Λ such that $(\lambda, \Lambda, \beta) = \mathscr{A}(r_m)$. Now, as an element of $S \operatorname{inv} T$, $r = \overline{r(\beta)}$ and (α, Γ, β) is isomorphic to the product in Proposition 4.4. Thus, by that proposition, $(\alpha, \Gamma, \beta) = \mathscr{A}(\varepsilon(A)r)$. This completes the proof of Theorem 4.5.

Our final canonical form is again graphical. We will use the following concept.

Let $\mathscr{T}$ be a finite rooted tree, with root α, say. Then $\mathscr{T}$ is implicitly directed "away from" α. We call $\mathscr{T}$ a *bilabelled tree over* $S \cup T$ if its vertices are labelled by idempotents of $S \cup T$ and its edges are labelled by nonidempotents of $S \cup T$ so that:

(i) if an edge is labelled by u then its initial vertex is labelled by uu^{-1} and its terminal vertex is labelled from the alternate factor S or T;

(ii) edges with the same initial vertex have different labels.

In such a tree, successive edges are labelled alternately from S and from T; for any vertex λ other than possibly the root, the edges leaving λ are labelled from the $\mathscr{R}$-class of the label of λ; the root may be labelled twice, in which case the outgoing edges are labelled from one of the two $\mathscr{R}$-classes of the labels of α. (To avoid this complication it is sometimes convenient to split $\mathscr{T}$ into two components, one whose outgoing edges from the root are labelled from S and another whose corresponding edges are labelled from T, see [**3**].)

Let λ be a vertex of such a tree. There is a unique simple path to λ from α. Let $v(\lambda) = v_1 \ldots v_n$, where $v_1, \ldots, v_{n-1}$ label the sequence of edges of the path and v_n labels λ itself. Then $v(\lambda)$ is a left reduced word in $S \operatorname{sgp} T$. Put $A(\mathscr{T}) = \{v(\lambda) : \lambda \in V(\mathscr{T})\}$, where $V(\mathscr{T})$ is the vertex set of $\mathscr{T}$. Then it is easily verified that $A(\mathscr{T})$ is a canonical set for $S \operatorname{inv} T$.

An *associate* of a bilabelled tree is either 1 or a reduced word $r = r_1 \ldots r_m$ such that $r_1 \ldots (r_m r_m^{-1}) \in A(\mathscr{T})$ (that is, $r_1, \ldots, r_{m-1}$ label the edges of a simple path in $\mathscr{T}$ and r_m belongs to the $\mathscr{R}$-class of the label of the path's terminal vertex).

To associate a bilabelled tree with an element of $S \operatorname{inv} T$, we again select presentations $(X \mid P)$ and $(Y \mid Q)$ for S and T, respectively. Let $\mathscr{A} = (\alpha, \Gamma, \beta)$ be an automaton of the form described in Theorem 4.1. The *skeleton* $\mathrm{sk}(\mathscr{A})$ of $\mathscr{A}$ is constructed as follows. The vertices of $\mathrm{sk}(\mathscr{A})$ are the nodes of $\mathscr{A}$ (relative to the root α). If the vertex λ is the root of the lobe Λ, label it by the element of S or of T represented by the automaton $(\lambda, \Lambda, \lambda)$. (The root α may be labelled from both S and T if it is an intersection point.) For each remaining node γ of Λ, if any, there is an edge from λ to γ labelled by the element of S or of T represented by the automaton $(\lambda, \Lambda, \gamma)$.

Then $\mathrm{sk}(\mathscr{A})$ is a bilabelled tree over $S \cup T$. Moreover, $A(\mathscr{A}) = A(\mathscr{T})$, as is clear from a comparison of the definitions of the two sets, and $r(\mathscr{A})$ is an associate of $\mathrm{sk}(\mathscr{A})$. The map $\mathscr{A} \to (A(\mathscr{A}), r(\mathscr{A}))$, therefore, factors as

$$\mathscr{A} \to (\mathrm{sk}(\mathscr{A}), r(\mathscr{A})) \to (A(\mathrm{sk}(\mathscr{A})), r(\mathscr{A})),$$

and an application of Theorem 4.5 completes the proof of the following theorem.

THEOREM 4.7. *The map* $\mathscr{A} \to (\mathrm{sk}(\mathscr{A}), r(\mathscr{A}))$ *is a* 1–1 *correspondence from the set of isomorphism classes of Schützenberger automata for* $S \operatorname{inv} T$, *with respect to* $\operatorname{Inv}\langle X \cup Y \mid P \cup Q \rangle$, *to the set of pairs* $(\mathscr{T}, a)$, *where* $\mathscr{T}$ *is a bilabelled tree over* $S \operatorname{inv} T$ *and* a *is an associate of* $\mathscr{T}$.

COROLLARY 4.8. *The elements of* $S \operatorname{inv} T$ *are uniquely representable in the form* $\varepsilon(A(\mathscr{T}))a$, *where* $\mathscr{T}$ *is a bilabelled tree over* $S \operatorname{inv} T$ *and* a *is an associate of* $\mathscr{T}$.

That these canonical forms are independent of the chosen presentations follows from the corresponding fact for canonical pairs.

A slight variant of this last description was used by Jones [3] to describe the free product of two E-unitary inverse semigroups. Margolis and Meakin [7] described the free product of two E-unitary inverse monoids, in the category of inverse monoids, in a graphical form, using the Cayley graph of the free group product of the maximal group homomorphic images of the inverse monoids. This form is strongly related to Corollary 4.8. For further details, see the cited papers.

5. Applications. By using Theorem 4.1, all the properties of free products deduced in [4] by the use of canonical pairs, may be re-derived by using Theorems 3.4–3.8 of [12] and related results. For instance, suppose $w_1, w_2 \in (X \cup X^{-1} \cup Y \cup Y^{-1})^+$ and $\mathscr{A}(w_1) = (\alpha_1, \Gamma_1, \beta_1)$, $\mathscr{A}(w_2) = (\alpha_2, \Gamma_2, \beta_2)$. Then $w_1\tau\mathscr{D}w_2\tau$ if and only if $\Gamma_1 \cong \Gamma_2$; $w_1\tau\mathscr{R}w_2\tau$ if and only if $\Gamma_1 \cong \Gamma_2$ and $\alpha_1 = \alpha_2$. This leads to the following result.

PROPOSITION 5.1. *If every* $\mathscr{R}$*-class of* S *and every* $\mathscr{R}$*-class of* T *is finite, then every* $\mathscr{R}$*-class of* $S \operatorname{inv} T$ *is finite.*

Proof. It is immediate from the definition of the Schützenberger graph itself that the hypothesis is equivalent to the finiteness of every Schützenberger graph of S and of T. By Theorem 4.1, every Schützenberger graph of $S \operatorname{inv} T$ has finitely many lobes, each of which is a Schützenberger graph for either S or T. Hence every Schützenberger graph for $S \operatorname{inv} T$ is finite and thus every $\mathscr{R}$-class is finite.

We will use this proposition to prove the following new result.

THEOREM 5.2. *The free product of two residually finite inverse semigroups is again residually finite.*

Proof. We first show that the semigroups may be assumed finite. This is most easily accomplished in the framework of Jones's canonical forms. Let S and T be residually finite inverse semigroups. Thus for any finite subset F of S there exists a homomorphism of S into a finite inverse semigroup that separates the members of F, and similarly for T.

Let $v_1, v_2 \in S \operatorname{inv} T$, $v_1 \neq v_2$. By Theorem 4.5, there are canonical sets A_1, A_2 in $S \operatorname{sgp} T$ and associates a_1, a_2 respectively, such that $v_1 = \varepsilon(A_1)a_1$ and $v_2 = \varepsilon(A_2)a_2$ in $S \operatorname{inv} T$. Let Z_0 be the set of elements of $S \cup T$ that appear as letters in words in

$A_1 \cup A_2 \cup \{a_1, a_2\}$ and let $Z = Z_0 \cup \{zz^{-1}: z \in Z_0\}$. (In terms of Schützenberger automata, assuming presentations for S and for T, we find that the set Z consists of the elements of $S \cup T$ represented by the individual Schützenberger automata (λ, Λ, μ), where Λ is a lobe of $\mathscr{A}(v_1)$ or of $\mathscr{A}(v_2)$ with root λ, and μ is any node of Λ. Alternatively, in the context of Theorem 4.7 and its corollary, Z consists of the union of the sets of vertex labels and of edge labels of the bilabelled trees corresponding to v_1 and v_2, together with the letters appearing in their associates.)

By assumption, there exist finite inverse semigroups S', T' and homomorphisms $\phi_S: S \to S'$, $\phi_T: T \to T'$, which separate the sets $Z \cap S$ and $Z \cap T$ respectively. These two homomorphisms jointly extend to a homomorphism $\phi: S \text{ sgp } T \to S' \text{ inv } T'$ and to a homomorphism $\bar{\phi}: S \text{ inv } T \to S' \text{ inv } T'$, by freeness. As an element of $S \text{ inv } T$, v_1 is mapped by $\bar{\phi}$ to $(\varepsilon(A_1)a_1)\bar{\phi} = \varepsilon(A_1\phi)a_1\phi$. Now, for each word $u = u_1 \ldots u_n \in A_1$, $u\phi = u_1\phi \ldots u_n\phi$ is again left reduced since, for each i, the homomorphism ϕ_S or ϕ_T, as appropriate, separates $u_i\phi$ and $u_iu_i^{-1}(\phi)$. The set $A_1\phi$ is clearly prefix-closed, because A_1 is; and uniqueness of last letters for $A_1\phi$ follows from that for A_1 because ϕ is bijective from A_1 to $A_1\phi$. Hence $A_1\phi$ is a canonical set for S'. Since a_1 is an associate for A_1, $a_1\phi$ is an associate for $A_1\phi$. Thus the representation $(\varepsilon(A_1)a_1)\bar{\phi} = \varepsilon(A_1\phi)a_1\phi$ for $v_1\bar{\phi}$ is the canonical one in $S' \text{ inv } T'$. Similar results for v_2 apply *mutatis mutandi.*

Since $v_1 \neq v_2$, $(A_1, a_1) \neq (A_2, a_2)$. It follows from bijectivity of ϕ on Z that $(A_1\phi, a_1\phi) \neq (A_2\phi, a_2\phi)$. Hence $v_1\bar{\phi} \neq v_2\bar{\phi}$ in $S' \text{ inv } T'$.

It is therefore sufficient to separate $v_1\bar{\phi}$ and $v_2\bar{\phi}$ in some finite homomorphic image of $S' \text{ inv } T'$. But, by Lemma 5.1, since S' and T' are finite, $S' \text{ inv } T'$ has finite $\mathscr{R}$-classes. The proof is completed by the following general result.

LEMMA 5.3. *If an inverse semigroup has finite $\mathscr{R}$-classes then it is residually finite.*

Proof. This is an immediate consequence of the faithfulness of the direct sum of the Schützenberger representations of any inverse semigroup S [**1**, Theorem 3.21], thereby embedding S in a direct product of semigroups of row-monomial matrices, each of finite dimension over a finite subgroup of S. (See [**1**, Sections 3.5, 3.6].)

An alternative proof of the lemma uses the Schützenberger automata of S, with respect to some presentation $(X \mid P)$. Given $u, v \in (X \cup X^{-1})^+$ such that $u\tau \neq v\tau$, without loss of generality $u\tau > v\tau$; then there is a homomorphism of S into the syntactic monoid of the finite Schützenberger automaton $\mathscr{A}(u)$ that separates $u\tau$ and $v\tau$. We omit the details.

REFERENCES

1. A. H. Clifford and G. B. Preston, *The algebraic theory of semigroups*, vol. I (American Mathematical Society, 1961).

2. J. E. Hopcroft and J. D. Ullman, *Introduction to automata theory, languages and computation* (Addison Wesley, 1979).

3. P. R. Jones, A graphical representation for the free product of E-unitary inverse semigroups, *Semigroup Forum* **24** (1982), 195–221.

4. P. R. Jones, Free products of inverse semigroups, *Trans. Amer. Math. Soc.* **282** (1984), 293–317.

5. N. Knox, *On the inverse semigroup coproduct of an arbitrary collection of groups* (Ph.D. thesis, Tennessee State University, 1974).

6. D. B. McAlister, Inverse semigroups generated by a pair of subgroups, *Proc. Roy. Soc. Edinburgh Sect. A* **77** (1976), 9–22.

7. S. W. Margolis and J. Meakin, E-unitary inverse monoids and the Cayley graph of a group presentation, *J. Pure Appl. Algebra* **58** (1989), 45–76.

8. J. Meakin, Automata and the word problem, in J. E. Pin (ed.), *Formal properties of finite automata and applications*, Lecture Notes in Comput. Sci. **386** (1989), 89–103.

9. W. D. Munn, Free inverse semigroups, *Proc. London Math. Soc.* (3) **29** (1974), 385–404.

10. M. Petrich, *Inverse semigroups* (Wiley, 1984).

11. J. R. Stallings, Topology of finite graphs, *Invent. Math.* **71** (1983), 551–565.

12. J. B. Stephen, Presentations of inverse monoids, *J. Pure Appl. Algebra* **63** (1990), 81–112.

13. J. B. Stephen, *Applications of automata theory to presentations of monoids and inverse monoids* (Ph.D. thesis, University of Nebraska, 1987).

PETER R. JONES
DEPARTMENT OF MATHEMATICS
STATISTICS AND COMPUTER SCIENCE
MARQUETTE UNIVERSITY
MILWAUKEE WI 53233

STUART W. MARGOLIS
DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF NEBRASKA
LINCOLN NE 68588

JOHN MEAKIN
DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF NEBRASKA
LINCOLN NE 68588

JOSEPH B. STEPHEN
DEPARTMENT OF MATHEMATICAL SCIENCES
NORTHERN ILLINOIS UNIVERSITY
DEKALB IL 60115