Lecture 6f: $FO^2(<)$, Δ_2 and **DA**

In the previous lecture we showed that the set of languages definable in Σ_1 is precisely the set of languages that are upward closed (w.r.t. substrings). This immediately means that the set of languages that are definable in Π_1 are precisely those that are downward closed. What about languages that are definable in Σ_1 and Π_1? It is easy to check that a language is both and upwards and downwards closed if and only if it is Σ^* or $\emptyset$ (if it contains any string, it contains ϵ and so contains Σ^*).

In general, we write Δ_i for the set of languages in $\Sigma_i \cap \Pi_i$. So, what we have just shown is that Δ_1 is the uninteresting class of languages $\{\emptyset, \Sigma^*\}$. In general, the class Δ_i is interesting as it is the largest complementation closed class within Σ_i (or Π_i) and hence the largest class closed under all boolean operations ($\cup$,$\cap$ and complement) within Σ_i (or Π_i). The class Δ_2 has very many interesting characterizations and this is the topic of this lecture.

Recall that we characterized Σ_2 using the class of ordered monoids satisfying $ese \leqslant e$ for each idempotent e and $s \in S_e$. The following is a direct consequence of this result.

Proposition 1 *A language L is in Δ_2 if and only if it is recognised by a monoid that satisfies the identity $ese = e$ for each idempotent e and $s \in S_e$.*

Proof: Let $L \in \Delta_2$. Since it is in Σ_2, the ordered syntactic monoid of L, $\mathsf{oSyn}(L)$ satisfies $ese \leqslant_L e$ for all idempotents e and $s \in S_e$. Further the ordered syntactic monoid of $\overline{L}$ is just the syntactic monoid of L with the order $\leqslant_L^R$. Since $\overline{L}$ is also in Σ_2 we have $ese \leqslant_L^R e$ for all idempotents e and $s \in S_e$, i.e. $e \leqslant_L ese$ for all idempotents e and $s \in S_e$. Therefore we have $ese = e$ in the syntactic monoid of L for all idempteents e and $s \in S_e$.

For the converse, suppose L is recognized by a monoid $(M, ., 1)$ that satisfies $ese = e$ for all idempotents e and $s \in S_e$. Then L is in Σ_2 as the ordered monoid $(M, ., 1, =)$ recognizes L and satisfies $ese = e$ for all idempotents e and $s \in S_e$. Since the same (ordered) monoid also recognizes $\overline{L}$, L is also in in Π_2. ■

We write E_M for the set of idempotents in M. Next we show that the requirement in the previous proposition can be weakened to $ese = e$ for all idempotents e and $s \in J_{\geqslant}(e)$, where $J_{\geqslant}(e) = \{s \mid e \leqslant_J s\}$. One direction is trivial since $J_{\geqslant}(e) \subseteq S_e$, while the other direction requires some work. We now derive an important property of the regular $\mathcal{D}$-classes of the monoids that satisfy $ese = e$ for all $e \in E_m$, $s \in J_{\geqslant}(e)$.

Lemma 2 *Let $(M, ., 1)$ be a monoid that satsifies $ese = e$ for all $e \in E_M$ and $s \in J_{\geqslant}(e)$. Then,*

1. *Every regular $\mathcal{H}$-class of M is trivial.*

2. *Every element of any regular $\mathcal{D}$-class is an idempotent.*

Thus, every regular $\mathcal{D}$-class is an idempotent (and hence aperiodic) semigroup.

Proof: Let H be a regular $\mathcal{H}$-class with an idempotent e and let $s \in H$. Then, $ese = e$, by the hypothesis. Moreover, $ese = s$, as H is a group with e as the identity. Thus $s = e$ and hence H is trivial.

Let s be in some regular $\mathcal{D}$-class D. Then there is an idempotent e in the $\mathcal{R}$ class of s and $es = s$. But, by the hypothesis, $ese = e$ and thus $s = es = eses = s^2$. Thus s is an idempotent.

If every element is an idempotent then, by the location lemma, D must be a semigroup as well. ■

An immediate consequence of this Lemma is the following:

Proposition 3 *A monoid M satisfies $ese = e$ for all $e \in E_M$ and $s \in S_e$ if and only if it satisifes $ese = e$ for all $e \in E_M$ and $s \in J_{\geqslant}(e)$.*

Proof: In one direction, the implication follows from the fact that $J_{\geqslant}(e) \subseteq S_e$. For the other direction, we first establish the following claim:

Claim: Let M satisfy $ese = e$ for all $e \in E_M$ and all $s \in J_{\geqslant}(e)$. For any s_1, s_2 such that $es_1e = e$ and $es_2e = e$ we have $es_1s_2e = e$.

Assuming the Claim, we can complete the proof of the proposition as follows. If $s \in S_e$, then by definition $s = s_1s_2 \ldots s_k$ where each $s_i \in J_{\geqslant}(e)$ and then by repeated application of the Claim, we have $es_1s_2 \ldots s_ke = e$ as required.

We now complete the proof by establishing the Claim. Let D be the $\mathcal{D}$-class of e. Since $es_1e = e$ we have $es_1\mathcal{R}e$ and similarly $s_2e\mathcal{L}e$. Applying Lemma 2, $(s_2e)(es_1) = s_2es_1$ is in the same $\mathcal{D}$-class as e and is an idempotent. But $s_2es_1\mathcal{L}es_1$ and $s_2es_1\mathcal{R}s_2e$ and thus, by the Location Lemma, es_1s_2e is in D and hence in the same $\mathcal{H}$-class as e. We then use Lemma 2 to conclude that $es_1s_2e = e$. The argument is summarized by the following egg-box diagram:

		$e = es_1s_2e$			es_1	
	$\cdots$		$\cdots$		$\cdots$	
		s_2e			s_2ees_1	

This completes the proof of the Claim. ■

Corollary 4 *A language is in Δ_2 if and only if it is recognized by a monoid satisfying $ese = e$ for all $e \in E_M$ and $s \in J_{\geqslant}(e)$.*

The class of monoids with the properties identified by Lemma 2 form a very important class:

Definition 5 *The class* **DA** *consists of all the monoids in which every regular* $\mathcal{D}$*-class is an aperiodic semigroup. We shall refer the class of languages recognized by monoids in* **DA** *by* **DA** *as well.*

It turns out that this class also charaterizes Δ_2.

Theorem 6 *A language is in* Δ_2 *iff it is in* **DA**.

Proof: As a consequence of Corollary 4 and Lemma 2 it follows that every language in Δ_2 is in **DA**. The converse requires some work:

Let D be the $\mathcal{D}$-class of an idempotent e. Since D is an aperiodic semigroup

1. Every $\mathcal{H}$-class in D is trivial. (Since $\mathcal{H}(e)$ is a group and D is aperiodic, $\mathcal{H}(e)$ is trivial. So all H-classes in D are trivial.)

2. Every element in D is an idempotent. (Since D is an aperiodic semigroup, this is implied by the Location Lemma.)

Let $e \leqslant_J s$. Then, $e = xsy$ and let $f = (syx)^N$ where N is the idempotent power of syx. Clearly $e \leqslant_J f$ and further, since $e = e^{N+1}$, $f \leqslant_J e$. Thus $e\mathcal{J}f$ and thus $e\mathcal{D}f$. This means $ef\mathcal{D}e$. But $ef \leqslant_R es$ and $es \leqslant_L e$.

■

References

[1] M. Bojanczyk: "Factorization Forests", *Proceedings of DLT 2009*, Springer LNCS 5583 (2009) 1-17.

[2] T. Colcombet: "Green's Relations and their Use in Automata Theory", *Proceedings of LATA 2011*, Spring LNCS 6638 (2011) 1-21.

[3] V Diekart, P Gastin and M Kufleitner: "A Survey on Small Fragments of First-Order Logic over Finite Words", International Journal of the Foundations of Computer Science **19(3)**, 2008.

[4] J.E.Pin: *Mathematical Foundations of Automata Theory*, MPRI Lecture Notes.