EF-Games over Linear Orders

Let $m \geq 1$ be a natural number, and $\mathcal{A} = (A, <^{\mathcal{A}})$ and $\mathcal{B} = (B, <^{\mathcal{B}})$ be linear orderings of lengths ℓ_A and ℓ_B, respectively.

Then $\mathcal{A} \equiv_m \mathcal{B}$ iff ($\ell_A = \ell_B$ or $\ell_A, \ell_B \geq 2^m - 1$).

Let's first show how Spoiler can distinguish structures of length less then $2^m - 1$. This is the implication from left to right (or, more precisely, the contrapositive thereof).

"$\Rightarrow$"
Assume $\ell_A \neq \ell_B$, say $\ell_A < \ell_B$, and one of ℓ_A or ℓ_B strictly smaller than $2^m - 1$. So $\ell_A < 2^m - 1$.

By Induction on m we show that Spoiler can win in m rounds.
For $m = 2$ we get $\ell_A < 2^2 - 1 = 4 - 1 = 3$, so $\ell_A \leq 2$. If either structure has strictly less than 2 elements, it's easy to win in 2 rounds. So the case remaining is a 2 element order versus an order with at least 3 Elements. Place a pebble in the larger(!) structure so that there is at least one element to either side. Duplicator has to play an "end-node" in the smaller structure.

In the degenerate case $m = 1$, we have that $\ell_A = 0$. Formally these are no legal first-order structures, as we require the universe to be non-empty. Nevertheless, the argument perfectly extends and, in fact, gives the more natural induction start. Spoiler just puts a pebble on an element of $\mathcal{B}$ and Duplicator can't match by a move in $\mathcal{A}$.

So let $m \geq 3$.
In case $\ell_B \geq 2^m - 1$ Spoiler puts a pebble "in the middle" of $\mathcal{B}$, such that on either side there are at least there are at least $2^{m-1} - 1$ pebbles left. Duplicator necessarily has to split $\mathcal{A}$ such that at least one part has length strictly less than $2^{m-1} - 1$. IH applies.
Otherwise $\ell_B < 2^m - 1$, so $\ell_B \leq 2^{m-1} - 2$, so $\ell_A \leq 2^{m-1} - 3$. In this case, Spoiler puts a pebble "in the middle" of $\mathcal{A}$, i.e., such that on either side there are at strictly less than $2^{m-1} - 1$ nodes left. Note that $2^{m-1} - 2 + 1 + 2^{m-1} - 2 = 2^{m-1} - 3$.
No matter how Duplicator plays, one of the parts will have different lengths, with the smaller part of size strictly less than $2^{m-1} - 1$.
IH applies.

For the other direction we have to provide a strategy for Duplicator to last m rounds. This is trivial in the case that $\ell_A = \ell_B$.

"$\Leftarrow$"
Assume $\ell_A, \ell_B \geq 2^m - 1$.

Roughly speaking, if Spoiler plays close to a border, we take the same distance from that border; otherwise, the structures will remain large, and hence indistinguishable.

Spoiler's move in, say, $\mathcal{A}$ will split up $\mathcal{A}$ in orders of length $\ell_{A,1}$ and $\ell_{A,2}$ with $\ell_{A,1} + \ell_{A,2} + 1 \ell_A \geq 2^m - 1$. So, at most one $\ell_{A,1}, \ell_{A,2}$ is smaller than $2^{m-1} - 1$, since $2^{m-1} - 1 + 2^{m-1} - 1 + 1 = 2^m - 1$.

If one, say $\ell_{A,1}$ is small, then split $\mathcal{B}$ into parts of size $\ell_{B,1} = \ell_{A,1}$ and $\ell_{B,2} \geq 2^{m-1} - 1$. Otherwise split $\mathcal{B}$ into parts both of size at least $2^{m-1} - 1$.

By Induction hypothesis, Duplicator can last another $m - 1$ rounds.

An important consequence of this observation is, that we cannot express the parity of the length of a word.

> **Corollary.** The property "the word has even length" is not first-order definable.

The argument is simply that we can find arbitrary long orders of different parity.

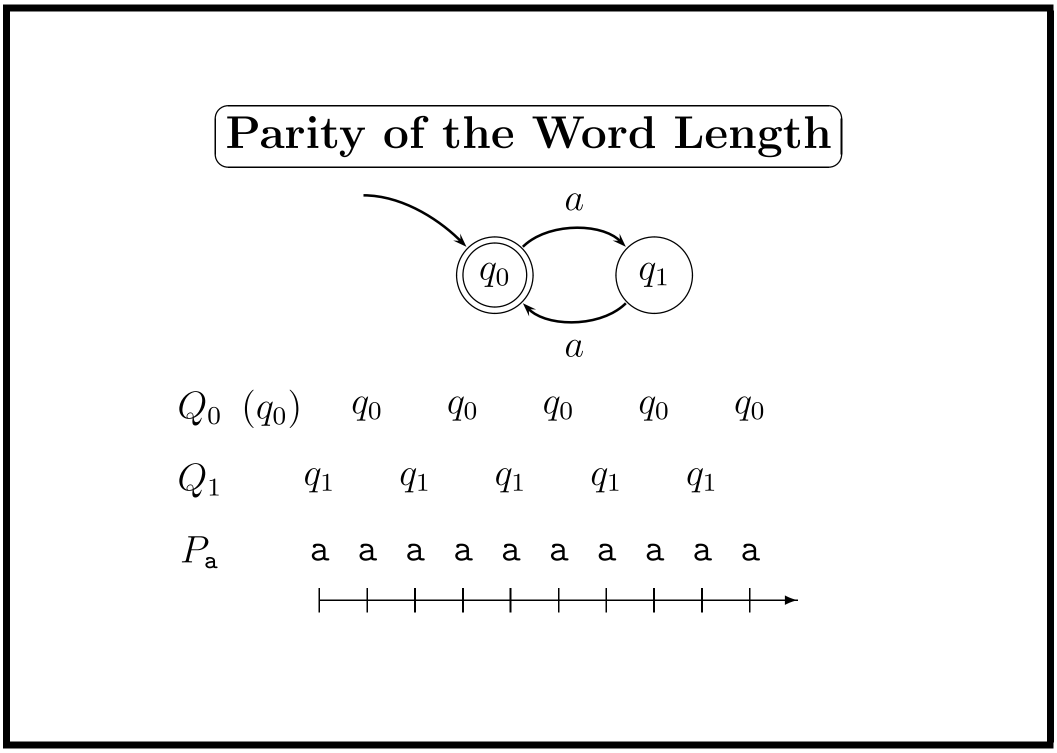
> **Proof.**
> Assume $\varphi \in \mathsf{FOL}[<]$ would define the property "even length". Let $m = \mathsf{qr}(\varphi)$. Then φ cannot distinguish $(\{1, 2, \ldots, 2^m\}, <)$ and $(\{1, 2, \ldots, 2^m + 1\}, <)$.

In particular, there are regular properties that are not first-order definable. We will later see, that all first-order definable properties in deed *are* regular.

2.5 Monadic Second-Order Logic

We have seen, that first-order logic cannot represent all regular properties; for example the property "has even length" cannot be expressed. Now let's try and add as little as possible to get all the regular properties.

Regular properties are given by finite automata. So, let's have a look at the one for "has even length".



This is almost a trace structure! The only new thing added are the two "custom predicates". So, as the new feature of first-order logic was, that we could talk freely about elements of the universe, we will now study a new logic where we can talk freely about unary properties.

As in the case of first-order logic, we presuppose an unbounded set of unary predicate variables

$$V^{(1)} = \{X_1, X_2, \ldots\}$$

Syntax of Monadic Second-Order Logic

The set $\mathsf{MSO}[R_1, \ldots, R_n]$ of first-order formulae over the relation symbols $R_1, \ldots, R_n$ is freely generated as follows.

- $(x = y)$ for variables x, y
- $R_j x_{i_1} \ldots x_{i_{n_j}}$ if R_j is of arity n_j
- $\boxed{Xx \text{ for } X \text{ predicate variable}}$
- $\top$, $\bot$, $\neg\phi$, $(\phi \wedge \psi)$, $(\phi \vee \psi)$ for formulae ϕ, ψ
- $\forall x\varphi$ and $\exists x\varphi$ for φ a formula and x a variable
- $\boxed{\forall X\varphi \text{ and } \exists X\varphi \text{ for } \varphi \text{ a fla and } X \text{ predicate var}}$

Again, we think of the second-order variables X as bound in $\forall X\varphi$ and $\exists X\varphi$. Correspondingly, we extend our definition of the set of free variables by adding two new clauses.

$$\begin{array}{lcl} \mathsf{fv}(Xx) & = & \{X, x\} \\ \mathsf{fv}(\forall X\varphi) = \mathsf{fv}(\exists X\varphi) & = & \mathsf{fv}(\varphi) \setminus \{X\} \end{array}$$

Just as we had an assignment $\eta\colon V \to A$ to have a meaning for the free first-order variables, we now also need a function $H\colon V^{(1)} \to \mathfrak{P}(A)$ to explain the meaning of the free second-order variables.

Semantics of Monadic Second-Order Logic

Let $\mathcal{A} = (A, R_1^{\mathcal{A}}, \ldots)$ and $\eta\colon V \to A$, and $H\colon V^{(1)} \to \mathfrak{P}(A)$.
Define $\mathcal{A}, H, \eta \models \varphi$ for $\varphi \in \mathsf{MSO}[R_1, \ldots]$ inductively.

- $\mathcal{A}, H, \eta \models x = y$ iff $\eta(x) = \eta(y)$
  $\mathcal{A}, H, \eta \models R_1 y_1 \ldots y_n$ iff $(\eta(y_1), \ldots, \eta(y_n)) \in R_i^{\mathcal{A}}$
  $\boxed{\mathcal{A}, H, \eta \models Xx \text{ iff } \eta(x) \in H(X)}$
- $\mathcal{A}, H, \eta \models \varphi \wedge \psi$ iff $\ldots$
  $\mathcal{A}, H, \eta \models \forall x \varphi$ iff $\mathcal{A}, H, \eta_x^a \models \varphi$ for all $a \in A$
- $\boxed{\begin{array}{l}\mathcal{A}, H, \eta \models \forall X \varphi \text{ iff } \mathcal{A}, H_X^U, \eta \models \varphi \text{ for all } U \in \mathfrak{P}(A)\\ \mathcal{A}, H, \eta \models \exists X \varphi \text{ iff } \mathcal{A}, H_X^U, \eta \models \varphi \text{ for some } U \in \mathfrak{P}(A)\end{array}}$

Again, we have a coincidence Lemma.

> **Lemma.**
> η, η', H, H' such that $\eta(x) = \eta'(x)$ for all $x \in \mathsf{fv}(\varphi)$, and $H(X) = H'(X)$ for all $X \in \mathsf{fv}(\varphi)$. Then $\mathcal{A}, \eta, H \models \varphi$ iff $\mathcal{A}, \eta', H' \models \varphi$.

This observation justifies to use the notation

$$\mathcal{A}, U_1, \ldots, U_k, a_1, \ldots, a_\ell \models \varphi$$

If $\mathsf{fv}(\varphi) \subset \{X_1, \ldots, X_k, x_1, \ldots, x_k\}$ with the order understood.