Memo on Logics over Finite Trees

Sylvain Schmitz
ENS Cachan, Université Paris-Saclay, France

We recall the syntax and semantics of two logics on finite trees: monadic second-order logic (MSO) and propositional dynamic logic (PDL). These are actually special cases of the same logics on finite relational structures, and we present the general framework.

1 Trees as Relational Structures

Relational Structures. We consider finite *relational signatures* $\sigma = ((R_i)_{1\leq i\leq n})$ where each relation symbol R_i has a fixed arity $r_i > 0$. A *σ-structure* is a tuple $\mathfrak{M} = (|\mathfrak{M}|, (R_i^M)_{1\leq i\leq n})$ where $|\mathfrak{M}|$ is the domain and each $R_i^{\mathfrak{M}}$ is an 'interpretation' of R_i as a relation in $|\mathfrak{M}|^{r_i}$; when the particular structure is clear from the context, we omit the $\mathfrak{M}$ superscripts in interpretation. A structure is finite if $|\mathfrak{M}|$ is finite.

Ranked Trees. Recall that a (finite ordered) ranked tree t over some finite ranked alphabet $\mathcal{F}$ can be seen as a partial function from $\mathbb{N}_{>0}$ to $\mathcal{F}$. Let $k \stackrel{\text{def}}{=} \max_{\mathcal{F}_i\neq\emptyset} i$ be the maximal arity in $\mathcal{F}$. We consider a finite set of atomic predicates A; typically $A = \mathcal{F}$, but in some applications one prefers $2^A = \mathcal{F}$. We shall use $A = \mathcal{F}$ here.

Ranked trees t in $T(\mathcal{F})$ can be seen as relational structures with domain $\mathrm{Pos}(t)$ over the signature $(\downarrow_1, \ldots, \downarrow_k, (P_f)_{f\in\mathcal{F}})$: we interpret the relations by

$$\downarrow_i \stackrel{\text{def}}{=} \{(p, pi) \in \mathrm{Pos}(t)^2\} \qquad \text{for all } 1 \leq i \leq k,$$
$$P_f \stackrel{\text{def}}{=} \{p \in \mathrm{Pos}(t) \mid t(p) = f\} \qquad \text{for all } f \in \mathcal{F}.$$

Other relational signatures are of course possible, for instance including

$$\downarrow \stackrel{\text{def}}{=} \{(p, pi) \in \mathrm{Pos}(t)^2 \mid i \in \mathbb{N}_{>0}\}\ ,$$
$$\downarrow^* \stackrel{\text{def}}{=} \{(p, pp') \in \mathrm{Pos}(t)^2 \mid p' \in \mathbb{N}_{>0}^*\}\ .$$

Unranked Trees. An unranked tree t over a finite alphabet Σ can similarly be seen as a relational structure with domain $\mathrm{Pos}(t)$ for the signature $(\downarrow, \rightarrow, (P_a)_{a\in\Sigma})$: we interpret the relations by

$$\downarrow \stackrel{\text{def}}{=} \{(p, pi) \in \mathrm{Pos}(t)^2 \mid i \in \mathbb{N}_{>0}\}\ ,$$
$$\rightarrow \stackrel{\text{def}}{=} \{(pi, p(i+1)) \in \mathrm{Pos}(t)^2 \mid i \in \mathbb{N}_{>0}\}\ ,$$
$$P_a \stackrel{\text{def}}{=} \{p \in \mathrm{Pos}(t) \mid t(p) = a\} \qquad \text{for all } a \in \Sigma.$$

Again, other relational signature are possible.

2 Monadic Second-Order Logic & Co.

Syntax. Consider a finite signature $\sigma = ((R_i)_{1\leq i\leq n})$. Let $\mathcal{X}_1$ and $\mathcal{X}_2$ be two infinite countable disjoint sets of first-order and second-order variables. The set of MSO(σ) formulæ is defined by the abstract syntax

$$\psi ::= R_i(x_1,\ldots,x_{r_i}) \mid x = x' \mid x \in X \mid \neg\psi \mid \psi \wedge \psi \mid \exists x.\psi \mid \exists X.\psi$$

where $1 \leq i \leq n$, $x, x', x_1, \dots \in \mathcal{X}_1$, and $X \in \mathcal{X}_2$. The set of FO(σ) formulæ is defined by removing second-order quantification and $x \in X$ predicates:

$$\psi ::= R_i(x_1,\ldots,x_{r_i}) \mid x = x' \mid \neg\psi \mid \psi \wedge \psi \mid \exists x.\psi\ .$$

Semantics. Given a σ-structure $\mathfrak{M} = (|\mathfrak{M}|, (R_i)_{1\leq i\leq n})$ and two valuations $\nu_1\colon \mathcal{X}_1 \to |\mathfrak{M}|$ and $\nu_2\colon \mathcal{X}_2 \to 2^{|\mathfrak{M}|}$, we say that $\mathfrak{M}$ *satisfies* ψ and write $\mathfrak{M} \models_{\nu_1,\nu_2} \psi$ in the following situations:

$$\begin{aligned}
&\mathfrak{M} \models_{\nu_1,\nu_2} R_i(x_1,\ldots,x_{r_i}) && \text{if } (\nu_1(x_1),\ldots,\nu_1(x_{r_i})) \in R_i\ ,\\
&\mathfrak{M} \models_{\nu_1,\nu_2} x = x' && \text{if } \nu_1(x) = \nu_1(x')\ ,\\
&\mathfrak{M} \models_{\nu_1,\nu_2} x \in X && \text{if } \nu_1(x) \in \nu_2(X)\ ,\\
&\mathfrak{M} \models_{\nu_1,\nu_2} \neg\psi && \text{if } \mathfrak{M} \not\models_{\nu_1,\nu_2} \psi\ ,\\
&\mathfrak{M} \models_{\nu_1,\nu_2} \psi \wedge \psi' && \text{if } \mathfrak{M} \models_{\nu_1,\nu_2} \psi \text{ and } \mathfrak{M} \models_{\nu_1,\nu_2} \psi'\ ,\\
&\mathfrak{M} \models_{\nu_1,\nu_2} \exists x.\psi && \text{if } \exists w \in |\mathfrak{M}|,\ \mathfrak{M} \models_{\nu_1[x\mapsto w],\nu_2} \psi\ ,\\
&\mathfrak{M} \models_{\nu_1,\nu_2} \exists X.\psi && \text{if } \exists S \subseteq |\mathfrak{M}|,\ \mathfrak{M} \models_{\nu_1,\nu_2[X\mapsto S]} \psi\ .
\end{aligned}$$

Examples on Unranked Trees. Over finite unranked trees and the signature $(\downarrow, \rightarrow, (P_a)_{a\in\Sigma})$, one typically defines the following first-order formulæ:

$$\mathsf{root}(x) \stackrel{\text{def}}{=} \neg\exists y(y \downarrow x) \qquad \mathsf{leaf}(x) \stackrel{\text{def}}{=} \neg\exists y(x \downarrow y)$$

$$\mathsf{first}(x) \stackrel{\text{def}}{=} \neg\exists y(y \rightarrow x) \qquad \mathsf{last}(x) \stackrel{\text{def}}{=} \neg\exists y(x \rightarrow y)$$

and the following MSO formulæ:

$$x \downarrow^* y \stackrel{\text{def}}{=} \forall X.(x \in X \wedge (\forall z \forall z'(z \in X \wedge z \downarrow z' \Rightarrow z' \in X)) \Rightarrow y \in X)$$

$$x \rightarrow^* y \stackrel{\text{def}}{=} \forall X.(x \in X \wedge (\forall z \forall z'(z \in X \wedge z \rightarrow z' \Rightarrow z' \in X)) \Rightarrow y \in X)\ .$$

Finally, we say that a tree t satisfies ψ if there exist ν_1 and ν_2 such that $t \models_{\nu_1,\nu_2} \psi$, and we define the *language* of ψ as $L(\psi) \stackrel{\text{def}}{=} \{t \in T(\Sigma) \mid \exists \nu_1, \nu_2,\ t \models_{\nu_1,\nu_2} \psi\}$.

3 Propositional Dynamic Logic

Here we assume that all the relational symbols in $\sigma = ((R_i)_{1\leq i\leq n}, (P_p)_{p\in A})$ to be either binary for all $(R_i)_{1\leq i\leq n}$ or unary for all $(P_p)_{p\in A}$. The definitions can actually be extended to higher arities.

Syntax. There are two sorts of PDL formulæ: *node formulæ* hold in particular points of the structure (called 'worlds' in the modal logic literature), while *path formulæ* hold between points. We present here a version of PDL with *converse*

$$\varphi ::= \top \mid p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \langle\pi\rangle\varphi \;, \qquad \text{(node formulæ)}$$

$$\pi ::= R_i \mid \varphi? \mid \pi^{-1} \mid \pi;\pi \mid \pi+\pi \mid \pi^* \;, \qquad \text{(path formulæ)}$$

where p ranges over A and $1 \leq i \leq n$.

Semantics. A node formula φ is satisfied in a world $w \in |\mathfrak{M}|$ of a σ-structure $\mathfrak{M} = (|\mathfrak{M}|, (R_i)_{1\leq i\leq n}, (P_p)_{p\in A})$, denoted $\mathfrak{M}, w \models \varphi$, in the following situations:

$$\begin{aligned}
&\mathfrak{M}, w \models \top && \text{always},\\
&\mathfrak{M}, w \models p && \text{if } w \in P_p,\\
&\mathfrak{M}, w \models \neg\varphi && \text{if } \mathfrak{M}, w \not\models \varphi \;,\\
&\mathfrak{M}, w \models \varphi \wedge \varphi' && \text{if } \mathfrak{M}, w \models \varphi \text{ and } \mathfrak{M}, w \models \varphi' \;,\\
&\mathfrak{M}, w \models \langle\pi\rangle\varphi && \text{if } \exists w' \in |\mathfrak{M}|,\ \mathfrak{M}, w, w' \models \pi \text{ and } \mathfrak{M}, w' \models \varphi' \;.
\end{aligned}$$

Similarly, a path formula π is satisfied between two worlds w and w' of $\mathfrak{M}$, denoted $\mathfrak{M}, w, w' \models \pi$, in the following situations:

$$\begin{aligned}
&\mathfrak{M}, w, w' \models R_i && \text{if } (w, w') \in R_i \;,\\
&\mathfrak{M}, w, w' \models \varphi? && \text{if } w = w' \text{ and } \mathfrak{M}, w \models \varphi \;,\\
&\mathfrak{M}, w, w' \models \pi^{-1} && \text{if } \mathfrak{M}, w', w \models \pi \;,\\
&\mathfrak{M}, w, w' \models \pi;\pi' && \text{if } \exists w'' \in |\mathfrak{M}|,\ \mathfrak{M}, w, w'' \models \pi \text{ and } \mathfrak{M}, w'', w' \models \pi' \;,\\
&\mathfrak{M}, w, w' \models \pi + \pi' && \text{if } \mathfrak{M}, w, w' \models \pi \text{ or } \mathfrak{M}, w, w' \models \pi' \;,\\
&\mathfrak{M}, w, w' \models \pi^* && \text{if } \exists n \in \mathbb{N},\ \exists w_1 = w, w_2, \ldots, w_{n-1}, w_n = w' \in |\mathfrak{M}|,\ \forall 1 \leq j < n,\ \mathfrak{M}, w_j, w_{j+1} \models \pi \;.
\end{aligned}$$

Satisfaction Sets. Alternatively, we can define the semantics through *satisfaction sets*:

$$[\![\varphi]\!]_{\mathfrak{M}} \stackrel{\text{def}}{=} \{w \in |\mathfrak{M}| \mid \mathfrak{M}, w \models \varphi\} \qquad [\![\pi]\!]_{\mathfrak{M}} \stackrel{\text{def}}{=} \{(w, w') \in |\mathfrak{M}|^2 \mid \mathfrak{M}, w, w' \models \pi\} \;.$$

One obtains for instance

$$[\![\langle\pi\rangle\varphi]\!]_{\mathfrak{M}} = ([\![\pi]\!]_{\mathfrak{M}})^{-1}([\![\varphi]\!]_{\mathfrak{M}}) \;, \qquad [\![\pi^*]\!]_{\mathfrak{M}} = [\![\pi]\!]^*_{\mathfrak{M}} \;.$$

Box Modalities. Finally, let us mention that the dual of the 'diamond' $\langle\pi\rangle$ is the 'box' $[\pi]\varphi \stackrel{\text{def}}{=} \neg\langle\pi\rangle\neg\varphi$:

$$\mathfrak{M}, w \models [\pi]\varphi \text{ if } \forall w' \in |\mathfrak{M}|,\ \mathfrak{M}, w, w' \models \pi \text{ implies } \mathfrak{M}, w' \models \varphi \;.$$

Examples on Unranked Trees. Over finite unranked trees and the signature $(\downarrow, \rightarrow, (P_a)_{a\in\Sigma})$, one typically defines the following path formulæ

$$\uparrow \stackrel{\text{def}}{=} \downarrow^{-1} \qquad\qquad \leftarrow \stackrel{\text{def}}{=} \rightarrow^{-1}$$

and node formulæ

$$\mathsf{root} \stackrel{\text{def}}{=} [\uparrow]\bot \qquad\qquad \mathsf{leaf} \stackrel{\text{def}}{=} [\downarrow]\bot$$

$$\mathsf{first} \stackrel{\text{def}}{=} [\leftarrow]\bot \qquad\qquad \mathsf{last} \stackrel{\text{def}}{=} [\rightarrow]\bot$$

Finally, we say that a tree t *satisfies* φ, denoted $t \models \varphi$, if it satisfies it at the root, i.e. $\varphi, \varepsilon \models \varphi$. The *language* of φ is $L(\varphi) \stackrel{\text{def}}{=} \{t \in T(\Sigma) \mid t \models \varphi\}$ the set of trees that satisfy the formula.