

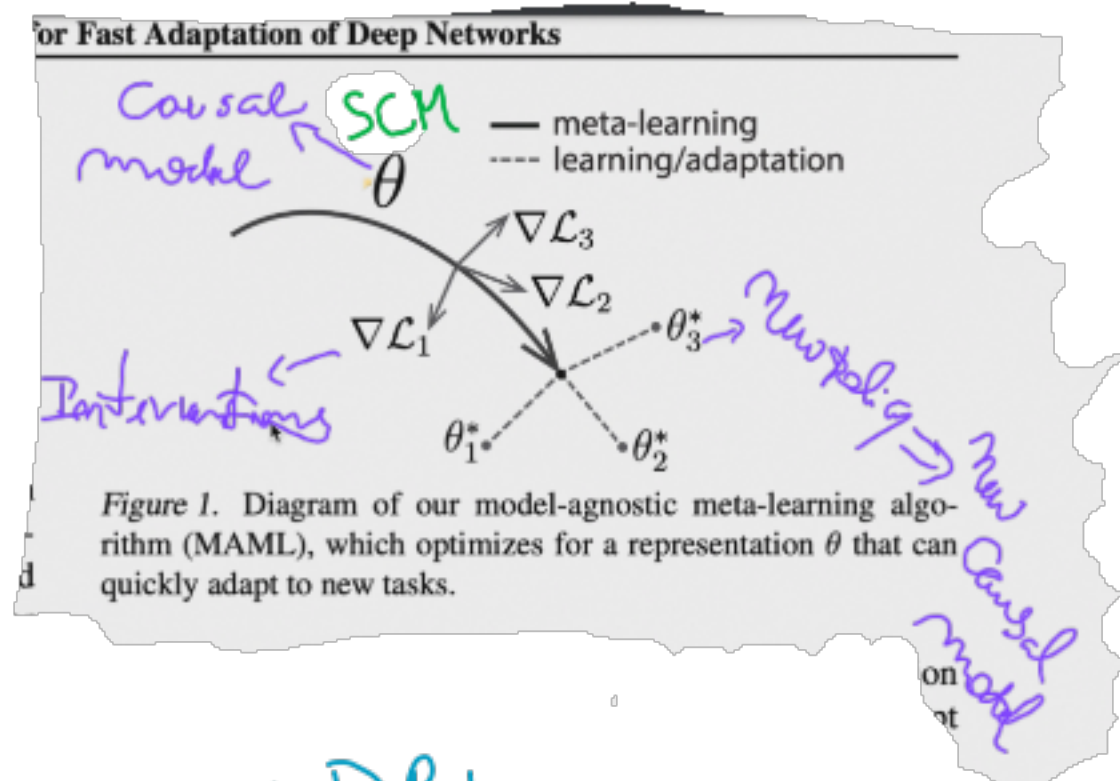
Environment Configuration

$\Phi_{1:N} = \langle \text{Desired Speed, Friction} \rangle$

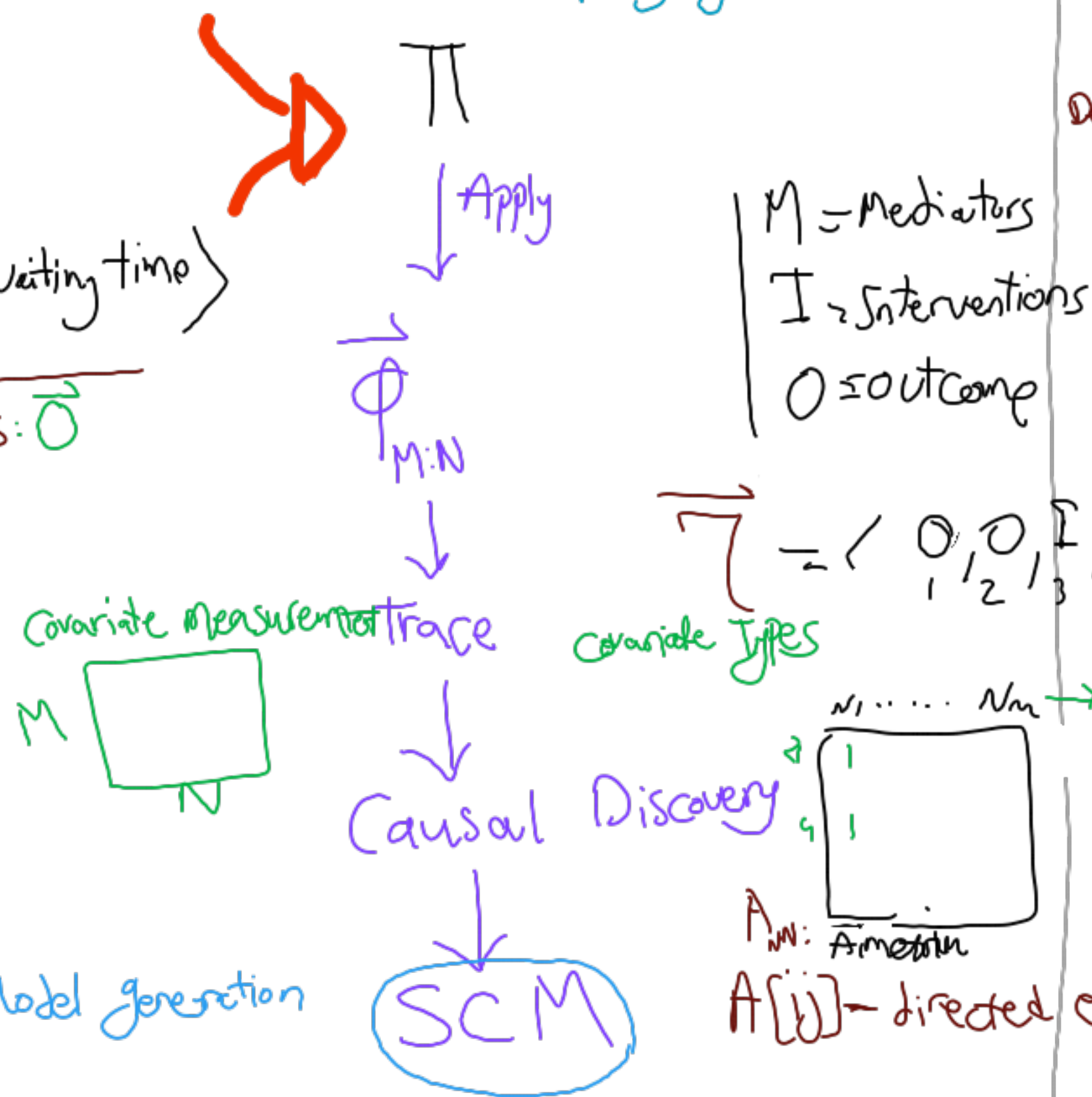
$R = \beta_x \langle \text{Collision, waiting time} \rangle$

outcomes: \vec{O}

Method: A2C

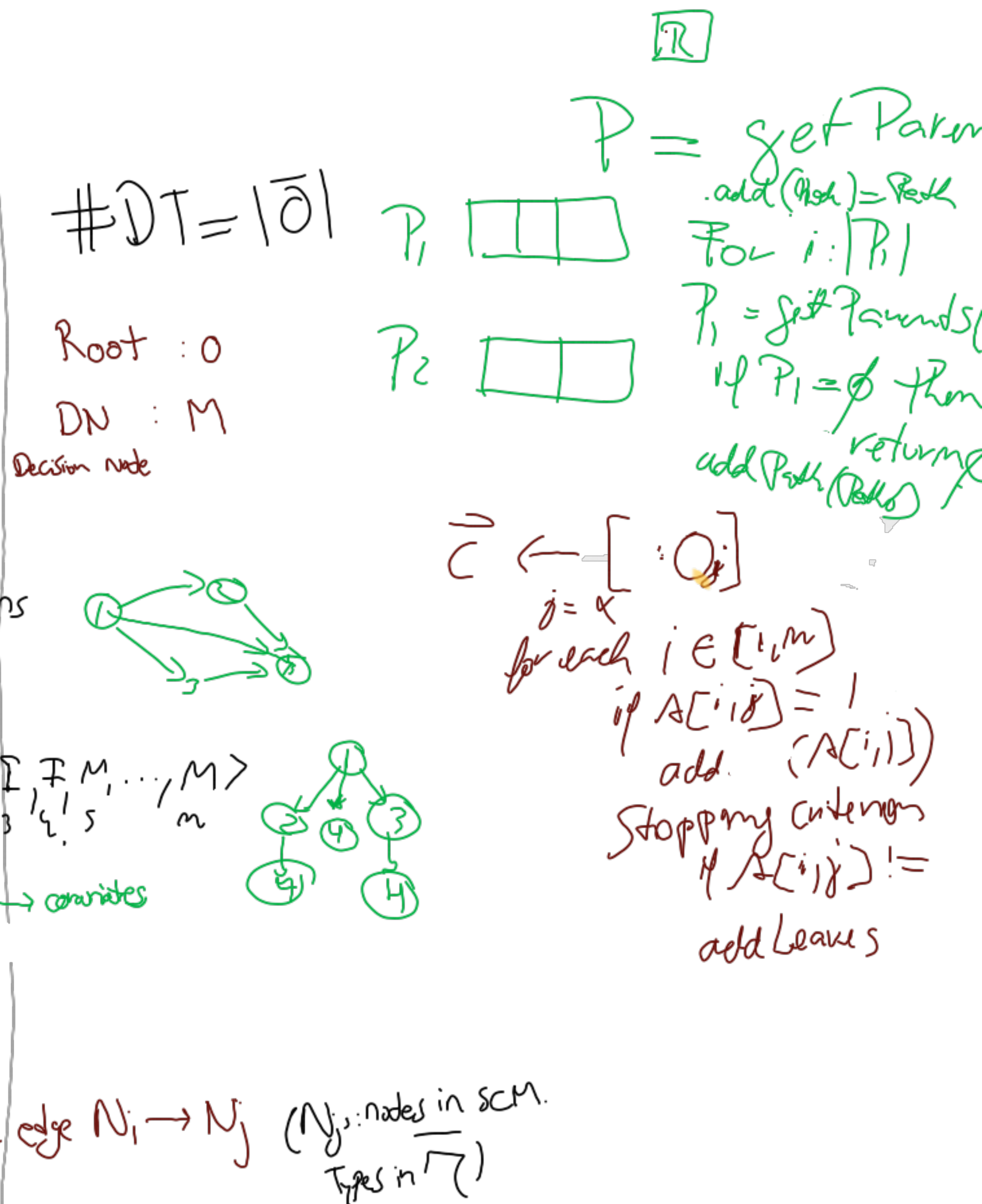


Step 1: DRL policy generation



Step 2: Causal Model generation

Step 3: Logical DT generation



nts (m)

$\{p_i\}$

ϕ