Merge-and-Shrink for SSPs with Prune Transformations — Proof Appendix

Thorsten Klößner[a], Álvaro Torralba[b], Marcel Steinmetz[c] and Silvan Sievers[d]

[a]Saarland University, Germany
[b]Aalborg University, Denmark
[c]LAAS-CNRS, Université de Toulouse, France
[d]University of Basel, Switzerland

Abstract

This proof appendix contains the sketched or omitted proofs of the ECAI 2024 paper "Merge-and-Shrink for SSPs with Pruning Transformations" by Klößner et al. The notation follows the original paper.

For the proofs below, recall again the definition of the transformed policy $tpol_{\tau,s}(\pi)$ as given in the paper:

Definition 6. *Let $\langle \Theta, \Theta', \sigma, \lambda \rangle$ be a transformation, let π be a policy for Θ and let $s \in S_\Theta$ be some starting state. The transformed policy $tpol_{\tau,s}(\pi)$ of π for s is defined by*

$$tpol_{\tau,s}(\pi)(p')(\mathfrak{T}') := \frac{\sum_{p \in tpath_\tau^{-1}(p')} \Pr_s^\pi[Cyl(p)] \cdot \sum_{\mathfrak{T} \in ttr_\tau^{-1}(\mathfrak{T}')} \pi(p)(\mathfrak{T})}{\sum_{p \in tpath_\tau^{-1}(p')} \Pr_s^\pi[Cyl(p)]}$$

if $\Pr_s^\pi[\bigcup_{p \in tpath_\tau^{-1}(p')} Cyl(p)] > 0$, and $tpol_{\tau,s}(\pi)(p')(\mathfrak{T}') := 0$ otherwise.

In the paper, the following theorem was stated without being formally shown. We will prove it in the following.

Theorem 2. *Let $\tau = \langle \Theta, \Theta', \sigma, \lambda \rangle \in$ **CONS$_{\mathbf{L+T}}$** and let $s \in S_\Theta$. For all policies π with $Reach^{\rightarrow}_{\Theta,\pi}(s) \subseteq dom(\sigma)$ and every set of target states $T \subseteq S_\Theta$:*

a) $Reach^{\rightarrow}_{\Theta,tpol_{\tau,s}(\pi)}(\sigma(s)) = \sigma(Reach^{\rightarrow}_{\Theta,\pi}(s))$ and
b) If $\pi \in Sols_\Theta(s,T)$, then $tpol_{\tau,s}(\pi) \in Sols_{\Theta'}(\sigma(s), \sigma(T))$.

To show this theorem, we first start with two lemmata. In the following, let $Prefs_\pi(s) := \{p \mid \Pr_s^\pi[Cyl(p)] > 0\}$ be the set of execution prefixes that can possibly be generated by a policy π with starting state s. Loosely speaking, the first lemma states that under the assumption **CONS$_\mathbf{L}$** and **CONS$_\mathbf{T}$**, as well as the assumption that every reachable state of the policy π with starting state s is kept by the transformation ($Reach^{\rightarrow}_{\Theta,\pi}(s) \subseteq dom(\sigma)$), every intermediate path p that can potentially be generated by π can be mapped to a transformed path $tpath_\tau(p)$ that exists within the transformed transition system Θ'. Additionally, if the original path reached a set of target states T, then the transformed path reaches the corresponding set of target states $\sigma(T)$.

Lemma 1. *Let $\tau = \langle \Theta, \Theta', \sigma, \lambda \rangle$ be a transformation, let π be a policy, let $s \in S_\Theta$ be a starting state and let $p \in FPaths(\Theta)$.*

*a) If $\tau \in$ **CONS$_{\mathbf{L+T}}$**, $Reach^{\rightarrow}_{\Theta,\pi}(s) \subseteq dom(\sigma)$ and $p \in Prefs_\pi(s)$, then $p \in dom(tpath_\tau)$ and $tpath_\tau(p) \in FPaths(\Theta')$.*
b) Let $T \subseteq S_\Theta$ be a set of target states. If the assumptions of Lemma 1a) hold and additionally $p \in Finish_\Theta(T)$, then $tpath_\tau(FPaths(\Theta')) \in Finish_{\Theta'}(T')$.

Proof. Let $p = s_0\mathfrak{T}_0 \dots s_n$ in the following.

We consider Lemma 1a) first. First of all, we have $s_0, \dots, s_n \in Reach^{\rightarrow}_{\Theta,\pi}(s)$ and also $supp(\mathfrak{T}_i) \subseteq Reach^{\rightarrow}_{\Theta,\pi}(s)$ for $0 \leq i < n$. Since $Reach^{\rightarrow}_{\Theta,\pi}(s) \subseteq dom(\sigma)$ by assumption and $\tau \in$ **CONS$_\mathbf{L}$**, we conclude that $s_0, \dots, s_n \in dom(\sigma)$ and $\mathfrak{T}_0, \dots, \mathfrak{T}_{n-1} \in dom(ttr_\tau)$. In particular, p is transformable, i.e., $p \in dom(tpath_\tau)$. Moreover, because of $\tau \in$ **CONS$_\mathbf{T}$**, we also have $ttr_\tau(\mathfrak{T}_0), \dots, ttr_\tau(\mathfrak{T}_{n-1}) \in T_{\Theta'}$ and therefore $tpath_\tau(p) \in FPaths(\Theta')$.

For Lemma 1b), we additionally have $p \in Finish_\Theta(T)$ so we have $s_n \in T$. As noted above, $s_n \in dom(\sigma)$, so we conclude $\sigma(s_n) \in \sigma(T)$ and $p \in Finish_{\Theta'}(\sigma(T))$ in particular. ☐

Next, we prove that, under the same assumptions as stated earlier, there is a nice relationship between the probability measures of the original policy with starting state s and the transformed policy with starting state $\sigma(s)$. Namely, the probability that the transformed policy generates an intermediate transformed path p' is equal to the probability that π generates *some* corresponding intermediate concrete path $p \in tpath_\tau^{-1}(p')$. The same also holds for the final path generated upon termination.

Lemma 2. *Let $\tau = \langle \Theta, \Theta', \sigma, \lambda \rangle$ be a transformation, let π be a policy and let $s \in S_\Theta$ be a starting state. Moreover, let $p' \in FPaths(\Theta')$ be a path of Θ'. If $\tau \in$ **CONS**$_{\mathbf{L+T}}$ and $Reach^{\rightarrow}_{\Theta,\pi}(s) \subseteq dom(\sigma)$, then:*

$$\Pr_{\sigma(s)}^{tpol_{\tau,s}(\pi)}[Cyl(p')] = \sum_{p \in tpath_\tau^{-1}(p')} \Pr_s^\pi[Cyl(p)] \tag{1}$$

$$\Pr_{\sigma(s)}^{tpol_{\tau,s}(\pi)}[p'] = \Pr_s^\pi\left[tpath_\tau^{-1}(p')\right] \tag{2}$$

Proof. Let $p' = s'_0 \mathfrak{T}'_0 \dots s'_n$ in the following.

Proof of Equation (1) First, we show Equation (1). By definition of the probability measure, we have

$$\Pr_{\sigma(s)}^{tpol_{\tau,s}(\pi)}[Cyl(p')] = [s'_0 = s'] \cdot \prod_{i=0}^{n-1} tpol_{\tau,s}(\pi)(s'_0\mathfrak{T}'_0 \dots s'_i)(\mathfrak{T}'_i) \cdot \delta_{\mathfrak{T}'_i}(s'_{i+1}).$$

If $\sum_{p \in tpath_\tau^{-1}(s'_0\mathfrak{T}'_0\dots s'_i)} \Pr_s^\pi[Cyl(p)] = 0$ for some prefix $s'_0\mathfrak{T}'_0 \dots s'_i$ of p', then for the right hand side of Equation (1) we have $\sum_{p \in tpath_\tau^{-1}(p')} \Pr_s^\pi[Cyl(p)] \le \sum_{p \in tpath_\tau^{-1}(s'_0\mathfrak{T}'_0\dots s'_i)} \Pr_s^\pi[Cyl(p)] = 0$ and the right hand side must be zero. The equation above has a factor $tpol_{\tau,s}(\pi)(s'_0\mathfrak{T}'_0 \dots s'_i) = 0$ in this case by definition of $tpol_{s,\pi}(\tau)$, thus the left hand side also evaluates to zero. All in all, both sides are zero under this assumption and the claim is shown.

Now assume $\sum_{p \in tpath_\tau^{-1}(s'_0\mathfrak{T}'_0\dots s'_i)} \Pr_s^\pi[Cyl(p)] > 0$. We apply the definition of $tpol_{\tau,s}(\pi)$ and obtain:

$$\Pr_{\sigma(s)}^{tpol_{\tau,s}(\pi)}[Cyl(p')] = [s'_0 = s'] \cdot \prod_{i=0}^{n-1} \frac{\sum\limits_{p \in tpath_\tau^{-1}(s'_0\mathfrak{T}'_0\dots s'_i)} \Pr_s^\pi[Cyl(p)] \cdot \sum\limits_{\mathfrak{T} \in ttr_\tau^{-1}(\mathfrak{T}'_i)} \pi(p)(\mathfrak{T}) \cdot \delta_{\mathfrak{T}'_i}(s'_{i+1})}{\sum\limits_{p \in tpath_\tau^{-1}(s'_0\mathfrak{T}'_0\dots s'_i)} \Pr_s^\pi[Cyl(p)]}$$

In the next step, we exploit that $\mathfrak{T} \in ttr_\tau^{-1}(\mathfrak{T}'_i)$ in the sum, from which it follows that $\delta_{\mathfrak{T}'_i} = lift[\sigma](\delta_\mathfrak{T})$.

$$\dots = [s'_0 = s'] \cdot \prod_{i=0}^{n-1} \frac{\sum\limits_{p \in tpath_\tau^{-1}(s'_0\mathfrak{T}'_0\dots s'_i)} \Pr_s^\pi[Cyl(p)] \cdot \sum\limits_{\mathfrak{T} \in ttr_\tau^{-1}(\mathfrak{T}'_i)} \pi(p)(\mathfrak{T}) \cdot lift[\sigma](\delta_\mathfrak{T})(s'_{i+1})}{\sum\limits_{p \in tpath_\tau^{-1}(s'_0\mathfrak{T}'_0\dots s'_i)} \Pr_s^\pi[Cyl(p)]}$$

We now apply the definition of $lift[\sigma]$, leading to:

$$\dots = [s'_0 = s'] \cdot \prod_{i=0}^{n-1} \frac{\sum\limits_{p \in tpath_\tau^{-1}(s'_0\mathfrak{T}'_0\dots s'_i)} \Pr_s^\pi[Cyl(p)] \cdot \sum\limits_{\mathfrak{T} \in ttr_\tau^{-1}(\mathfrak{T}'_i)} \pi(p)(\mathfrak{T}) \cdot \sum\limits_{s \in \sigma^{-1}(s'_{i+1})} \delta_\mathfrak{T}(s)}{\sum\limits_{p \in tpath_\tau^{-1}(s'_0\mathfrak{T}'_0\dots s'_i)} \Pr_s^\pi[Cyl(p)]}$$

$$\dots = [s'_0 = s'] \cdot \prod_{i=0}^{n-1} \frac{\sum\limits_{\langle p,\mathfrak{T},s\rangle \in tpath_\tau^{-1}(s'_0\mathfrak{T}'_0\dots s'_i) \times ttr_\tau^{-1}(\mathfrak{T}'_i) \times \sigma^{-1}(s'_{i+1})} \Pr_s^\pi[Cyl(p)] \cdot \pi(p)(\mathfrak{T}) \cdot \delta_\mathfrak{T}(s)}{\sum\limits_{p \in tpath_\tau^{-1}(s'_0\mathfrak{T}'_0\dots s'_i)} \Pr_s^\pi[Cyl(p)]}$$

Finally, we make use of the fact $\Pr_s^\pi[Cyl(p)] \cdot \pi(p)(\mathfrak{T}) \cdot \delta_\mathfrak{T}(s) = \Pr_s^\pi[Cyl(p\mathfrak{T}s)]$ and simplify:

$$\dots = [s'_0 = s'] \cdot \prod_{i=0}^{n-1} \frac{\sum\limits_{\langle p,\mathfrak{T},s\rangle \in tpath_\tau^{-1}(s'_0\mathfrak{T}'_0\dots s'_i) \times ttr_\tau^{-1}(\mathfrak{T}'_i) \times \sigma^{-1}(s'_{i+1})} \Pr_s^\pi[Cyl(p\mathfrak{T}s)]}{\sum\limits_{p \in tpath_\tau^{-1}(s'_0\mathfrak{T}'_0\dots s'_i)} \Pr_s^\pi[Cyl(p)]}$$

$$\dots = [s'_0 = s'] \cdot \prod_{i=0}^{n-1} \frac{\sum\limits_{p \in tpath_\tau^{-1}(s'_0\mathfrak{T}'_0\dots s'_i\mathfrak{T}'_i s_{i+1})'} \Pr_s^\pi[Cyl(p)]}{\sum\limits_{p \in tpath_\tau^{-1}(s'_0\mathfrak{T}'_0\dots s'_i)} \Pr_s^\pi[Cyl(p)]}$$

This is a telescoping product, in which every numerator cancels with the following denominator, and only the final numerator for $i = n-1$ remains. Note in particular that $[s'_0 = s'] = \Pr_s^\pi[\bigcup_{s_0 \in \sigma^{-1}(s'_0)} Cyl(s_0)]$. Therefore:

$$\dots = \sum_{p \in tpath_\tau^{-1}(p')} \Pr_s^\pi[Cyl(p)]$$

This concludes the proof of Equation (1).

Proof of Equation (2) To prove Equation (2), we first express the event $\{p'\}$ as the event of all executions with prefix p', without those which continue along some transition to some state. Then we simplify the equations using our assumptions and Equation (1).

$$\begin{aligned}
\Pr_{\sigma(s)}^{tpol_{\tau,s}(\pi)}[p'] &= \Pr_{\sigma(s)}^{tpol_{\tau,s}(\pi)}\left[Cyl(p') \setminus \Big(\bigcup_{\mathfrak{T}' \in T_{\Theta'}} \bigcup_{t' \in S_{\Theta'}} Cyl(p'\mathfrak{T}'t')\Big)\right] \\
&= \Pr_{\sigma(s)}^{tpol_{\tau,s}(\pi)}[Cyl(p')] - \sum_{\mathfrak{T}' \in T_{\Theta'}} \sum_{t' \in S_{\Theta'}} \Pr_{\sigma(s)}^{tpol_{\tau,s}(\pi)}[Cyl(p'\mathfrak{T}'t')] \\
&= \sum_{p \in tpath_\tau^{-1}(p')} \Pr_s^\pi[Cyl(p)] - \sum_{\mathfrak{T}' \in T_{\Theta'}} \sum_{t' \in S_{\Theta'}} \sum_{\substack{p \in tpath_\tau^{-1}(p') \\ \mathfrak{T} \in ttr_\tau^{-1}(\mathfrak{T}') \\ t \in \sigma^{-1}(t')}} \Pr_s^\pi[Cyl(p\mathfrak{T}t)] && \text{(by Equation (1))} \\
&= \sum_{p \in tpath_\tau^{-1}(p')} \Big(\Pr_s^\pi[Cyl(p)] - \sum_{\mathfrak{T} \in ttr_\tau^{-1}(T_{\Theta'})} \sum_{t \in \sigma^{-1}(S_{\Theta'})} \Pr_s^\pi[Cyl(p\mathfrak{T}t)]\Big) \\
&= \sum_{p \in tpath_\tau^{-1}(p')} \Big(\Pr_s^\pi[Cyl(p)] - \sum_{\mathfrak{T} \in T_{\Theta}} \sum_{t \in S_{\Theta}} \Pr_s^\pi[Cyl(p\mathfrak{T}t)]\Big) && \text{(by Lemma 1a))} \\
&= \sum_{p \in tpath_\tau^{-1}(p')} \Big(\Pr_s^\pi[Cyl(p)] - \Pr_s^\pi\Big[\bigcup_{\mathfrak{T} \in T_{\Theta}} \bigcup_{t \in S_{\Theta}} Cyl(p\mathfrak{T}t)\Big]\Big) \\
&= \sum_{p \in tpath_\tau^{-1}(p')} \Pr_s^\pi\left[Cyl(p) \setminus \Big(\bigcup_{\mathfrak{T} \in T_{\Theta}} \bigcup_{t \in S_{\Theta}} Cyl(p\mathfrak{T}t)\Big)\right] \\
&= \sum_{p \in tpath_\tau^{-1}(p')} \Pr_s^\pi[p] \\
&= \Pr_s^\pi[tpath_\tau^{-1}(p')]
\end{aligned}$$

This shows the claim. □

We now use both of these lemmata to prove the initial claim.

Theorem 2. *Let $\tau = \langle \Theta, \Theta', \sigma, \lambda \rangle \in$ **CONS$_{\mathbf{L+T}}$** and let $s \in S_\Theta$. For all policies π with $Reach^{\rightarrow}_{\Theta,\pi}(s) \subseteq dom(\sigma)$ and every set of target states $T \subseteq S_\Theta$:*

a) $Reach^{\rightarrow}_{\Theta, tpol_{\tau,s}(\pi)}(\sigma(s)) = \sigma(Reach^{\rightarrow}_{\Theta,\pi}(s))$ and
b) If $\pi \in Sols_\Theta(s,T)$, then $tpol_{\tau,s}(\pi) \in Sols_{\Theta'}(\sigma(s), \sigma(T))$.

Proof. The inclusion from left to right holds without any assumptions, since $\Pr_{\sigma(s)}^{tpol_{\tau,s}(\pi)}[Cyl(p')] > 0$ for some path $p' \in FPaths(\Theta')$ implies that there is a concrete path $p \in tpath_\tau^{-1}(p')$ with $\Pr_s^\pi[Cyl(p)] > 0$ by Equation (1), where we have $last(p) \in dom(\sigma)$ and $last(p') = \sigma(last(p))$ in particular.

For the other direction, acknowledge that for every path $p \in Prefs_\pi(s)$, we have $p \in dom(tpath_\tau)$ and $tpath_\tau(p) \in FPaths(\Theta')$ by Lemma 1a) in the context of our assumptions. Therefore,

$$0 < \Pr_s^\pi[Cyl(p)] \leq \sum_{p \in tpath_\tau^{-1}(tpath_\tau(p))} \Pr_s^\pi\left[Cyl(p)\right] = \Pr_{\sigma(s)}^{tpol_{\tau,s}(\pi)}[Cyl(tpath_\tau(p))]$$

where the last equality follows from Equation (1) and $\sigma(last(p)) = last(tpath_\tau(p))$ in particular.

For Theorem 2b), we use Equation (2), followed by Lemma 1b). In particular, from Lemma 1b) we obtain the fact that $tpath_\tau^{-1}(Finish_{\Theta'}(\sigma(T))) \supseteq Finish_\Theta(T) \cap Prefs_\pi(s)$ in the context of our assumptions. Hence:

$$\Pr_{\sigma(s)}^{tpol_{\tau,s}(\pi)}[Finish_{\Theta'}(\sigma(T))] = \Pr_s^\pi[tpath_\tau^{-1}(Finish_{\Theta'}(\sigma(T)))] \geq \Pr_s^\pi[Finish_\Theta(T) \cap Prefs_\pi(s)] = 1.$$

Thus, $tpol_{s,\pi}(\tau) \in Sols_{\Theta'}(\sigma(s), \sigma(T))$. □