A Theory of Merge-and-Shrink for Stochastic Shortest Path Problems
Technical Report

Thorsten Klößner[1], Álvaro Torralba[2], Marcel Steinmetz[1], Silvan Sievers[3]
[1]Saarland University, Germany,
[2]Aalborg University, Denmark,
[3]Basel University, Switzerland
{kloessner, steinmetz}@cs.uni-saarland.de, alto@cs.aau.dk, silvan.sievers@unibas.ch

This technical report contains the full versions of all sketched or omitted proofs of our ICAPS 2023 paper "A Theory of Merge-and-Shrink for Stochastic Shortest Path Problems". We follow the same notation as originally introduced in the paper.

Proof of Theorem 2

Recall that for a transformation $\tau = \langle \Theta, \Theta', \sigma, \lambda \rangle$, we defined the following two heuristics:

$$h^\tau_\Theta(s) := \begin{cases} J^*_{\Theta'}(\sigma(s)) & s \in \mathrm{dom}(\sigma) \\ \infty & s \notin \mathrm{dom}(\sigma) \end{cases} \qquad h^\tau_{\Theta'}(s') := \begin{cases} \max_{s \in \sigma^{-1}(s')} J^*_\Theta(s) & \sigma^{-1}(s') \neq \emptyset \\ 0 & \sigma^{-1}(s') = \emptyset \end{cases}$$

The statement of Theorem 2 was the following.

Theorem 2 *Let $\tau = \langle \Theta, \Theta', \sigma, \lambda \rangle$ be a transformation.*

(i) If τ is conservative, then h^τ_Θ is goal-aware, consistent, safe and admissible.
(ii) If τ is refinable, then $h^\tau_{\Theta'}$ is goal-aware, consistent, safe and admissible.
(iii) If τ is refinable, then h^τ_Θ is pessimistic.
(iv) If τ is exact, then h^τ_Θ is perfect.

As already mentioned in the paper, (iii) follows from (ii) since $h^\tau_\Theta(s) = J^*_{\Theta'}(\sigma(s)) \geq h^\tau_{\Theta'}(\sigma(s)) \geq J^*_\Theta(s)$ by admissibility if $s \in \mathrm{dom}(\sigma)$ and $h^\tau_\Theta(s) = \infty \geq J^*_\Theta(s)$ otherwise. Clearly, (iv) follows directly from (i) and (iii). Admissibility of the heuristics follows from goal-awareness, consistency and safety. To prove (i) and (ii), we split these statements into multiple parts. We first concentrate on goal-awareness and consistency only, and then leverage this intermediate result to also show safety and admissibility.

Theorem 2A *Let $\tau = \langle \Theta, \Theta', \sigma, \lambda \rangle$ be a transformation.*

(i) If τ is conservative, h^τ_Θ is goal-aware and consistent.
(ii) If τ is refinable, $h^\tau_{\Theta'}$ is goal-aware and consistent.

Proof. Let $\Theta = \langle S, L, c, T, G \rangle$ and $\Theta' = \langle S', L', c', T', G' \rangle$.

(i) Goal-Awareness Let $s \in G$. From $\mathbf{CONS_S}$ and $\mathbf{CONS_G}$ we can immediately conclude $\sigma(s) \in G'$ and therefore $h^\tau_\Theta(s) = J^*_{\Theta'}(\sigma(s)) = 0$.

(i) Consistency Let $s \in S$. We need to show $h^\tau_\Theta(s) \leq (\mathcal{B}_\Theta h^\tau_\Theta)(s)$. Assume $T(s) \neq \emptyset$, as otherwise this is trivial by definition of $\mathcal{B}_\Theta$. This also implies $T'(\sigma(s)) \neq \emptyset$ since $\mathrm{ind}_\tau(T(s)) \neq \emptyset$ by $\mathbf{CONS_S}$ and $\mathbf{CONS_L}$ and $\mathrm{ind}_\tau(T(s)) \subseteq T'(\sigma(s))$ by $\mathbf{CONS_T}$. Because $J^*_{\Theta'} \leq \mathcal{B}_{\Theta'} J^*_{\Theta'}$, we have by definition of h^τ_Θ and $\mathcal{B}_{\Theta'}$:

$$\begin{aligned} h^\tau_\Theta(s) &\leq \min_{\langle \sigma(s), \ell', \delta' \rangle \in T'(\sigma(s))} \Big[c'(\ell') + \sum_{t' \in S'} \delta'(t') J^*_{\Theta'}(t') \Big] \\ &\leq \min_{\langle \sigma(s), \ell', \delta' \rangle \in \mathrm{ind}_\tau(T(s))} \Big[c'(\ell') + \sum_{t' \in S'} \delta'(t') J^*_{\Theta'}(t') \Big] && \text{(by } \mathrm{ind}_\tau(T(s)) \subseteq T'(\sigma(s))\text{)} \end{aligned}$$

Next, acknowledge the simple mathematical fact that $\min_{x \in f(X)} g(x) = \min_{x \in X} g(f(x))$ for f total on X. In our case, $f = \text{ind}_\tau$ is total on $X = T(s)$ by $\mathbf{CONS_S}$ and $\mathbf{CONS_L}$. This allows us to minimize over $T(s)$ directly.

$$
\begin{aligned}
\ldots &= \min_{\langle s,\ell,\delta\rangle \in T(s)} \Big[c'(\lambda(\ell)) + \sum_{t' \in S'} \sigma_{\mathcal{D}}(\delta)(t') J^*_{\Theta'}(t')\Big] \\
&\leq \min_{\langle s,\ell,\delta\rangle \in T(s)} \Big[c(\ell) + \sum_{t' \in S'} \sigma_{\mathcal{D}}(\delta)(t') J^*_{\Theta'}(t')\Big] && \text{(by } \mathbf{CONS_C}\text{)} \\
&= \min_{\langle s,\ell,\delta\rangle \in T(s)} \Big[c(\ell) + \sum_{t' \in S'} \sum_{t \in \sigma^{-1}(t')} \delta(t) J^*_{\Theta'}(t')\Big] && \text{(by def. } \sigma_{\mathcal{D}}\text{)} \\
&= \min_{\langle s,\ell,\delta\rangle \in T(s)} \Big[c(\ell) + \sum_{t \in S} \delta(t) J^*_{\Theta}(\sigma(t))\Big] && \text{(by } \mathbf{CONS_S}\text{)} \\
&= \min_{\langle s,\ell,\delta\rangle \in T(s)} \Big[c(\ell) + \sum_{t \in S} \delta(t) h^\tau_{\Theta}(t)\Big] && \text{(by def. } h^\tau_\Theta\text{)} \\
&= (\mathcal{B}_\Theta h^\tau_\Theta)(s) && \text{(by def. } \mathcal{B}_\Theta\text{)}
\end{aligned}
$$

(ii) Goal-Awareness Let $s' \in G'$. If $\sigma^{-1}(s') = \emptyset$ we have $h^\tau_{\Theta'}(s') = 0$ by definition of h, otherwise $\mathbf{REF_G}$ guarantees $\sigma^{-1}(s') \subseteq G$ and $h^\tau_{\Theta'}(s') = 0$ since $J^*_\Theta(s) = 0$ for $s \in G$.

(ii) Consistency Let $s' \in S'$. We need to prove $h^\tau_{\Theta'}(s') \leq (\mathcal{B}_{\Theta'} h^\tau_{\Theta'})(s')$. Assume $T'(s') \neq \emptyset$, as otherwise this is trivial by definition of $\mathcal{B}_{\Theta'}$. We can also assume $\sigma^{-1}(s') \neq \emptyset$ as otherwise $h^\tau_{\Theta'}(s') = 0$ by definition of $h^\tau_{\Theta'}$ and the inequality is trivial since all terms are non-negative.

We choose an arbitrary $s \in \sigma^{-1}(s')$. Note that $T'(s') \subseteq \text{ind}_\tau(T(s))$ by $\mathbf{REF_T}$ and therefore $\emptyset \neq T(s)$ and $\text{ind}^{-1}_\tau(T'(s')) \subseteq T(s)$. Because $J^*_\Theta \leq \mathcal{B}_\Theta J^*_\Theta$, we have

$$
\begin{aligned}
J^*_\Theta(s) &\leq \min_{\langle s,\ell,\delta\rangle \in T(s)} \Big[c(\ell) + \sum_{t \in S} \delta(t) J^*_\Theta(t)\Big] \\
&\leq \min_{\langle s,\ell,\delta\rangle \in \text{ind}^{-1}_\tau(T'(s'))} \Big[c(\ell) + \sum_{t \in S} \delta(t) J^*_\Theta(t)\Big] && \text{(by } \text{ind}^{-1}_\tau(T'(s')) \subseteq T(s)\text{)}
\end{aligned}
$$

Now, notice that $\delta \in \text{dom}(\sigma_{\mathcal{D}})$ for $\langle s,\ell,\delta\rangle \in \text{ind}^{-1}_\tau(T'(s'))$. Thus, $\text{supp}(\delta) \subseteq \text{dom}(\sigma) = \sigma^{-1}(S')$ by definition of $\sigma_{\mathcal{D}}$.

$$
\begin{aligned}
\ldots &= \min_{\langle s,\ell,\delta\rangle \in \text{ind}^{-1}_\tau(T'(s'))} \Big[c(\ell) + \sum_{t \in \sigma^{-1}(S')} \delta(t) J^*_\Theta(t)\Big] && \text{(by } \text{supp}(\delta) \subseteq \sigma^{-1}(S')\text{)} \\
&= \min_{\langle s,\ell,\delta\rangle \in \text{ind}^{-1}_\tau(T'(s'))} \Big[c(\ell) + \sum_{\substack{t' \in S' \\ t \in \sigma^{-1}(t')}} \delta(t) J^*_\Theta(t)\Big] \\
&= \min_{\langle s,\ell,\delta\rangle \in \text{ind}^{-1}_\tau(T'(s'))} \Big[c'(\lambda(\ell)) + \sum_{\substack{t' \in S' \\ t \in \sigma^{-1}(t')}} \delta(t) J^*_\Theta(t)\Big] && \text{(by } \mathbf{REF_C}\text{)} \\
&\leq \min_{\langle s,\ell,\delta\rangle \in \text{ind}^{-1}_\tau(T'(s'))} \Big[c'(\lambda(\ell)) + \sum_{t' \in S'} \sigma_{\mathcal{D}}(\delta)(t') h^\tau_{\Theta'}(t')\Big] && \text{(by def. } \sigma_{\mathcal{D}} \text{ and } h^\tau_{\Theta'}\text{)}
\end{aligned}
$$

Lastly, we minimize over the transitions in $T'(s')$ directly instead.

$$
\begin{aligned}
\ldots &= \min_{\langle s',\ell',\delta'\rangle \in T'(s')} \Big[c'(\ell') + \sum_{t' \in S'} \delta'(t') h^\tau_{\Theta'}(t')\Big] \\
&= (\mathcal{B}_{\Theta'} h^\tau_{\Theta'})(s') && \text{(by def. } \mathcal{B}_{\Theta'}\text{)}
\end{aligned}
$$

All in all, $J^*_\Theta(s) \leq (\mathcal{B}_{\Theta'} h^\tau_{\Theta'})(s')$. Since $s \in \sigma^{-1}(s')$ was arbitrary, we have $h^\tau_{\Theta'}(s') = \max_{s \in \sigma^{-1}(s')} J^*_\Theta(s) \leq (\mathcal{B}_{\Theta'} h^\tau_{\Theta'})(s')$ by definition of $h^\tau_{\Theta'}$. $\square$

To show the remaining properties of Theorem 2 (i) and (ii), we apply Theorem 2A to a very specific type of transformation. To this end, for a PTS $\Theta = \langle S, L, c, T, G\rangle$, we introduce the *MaxProb compilation* of Θ as $\text{MP}(\Theta) := \langle S \cup \{s_G(\Theta)\}, L \cup \{\text{give\_up}(\Theta), \text{success}(\Theta)\}, c', T', \{s_G(\Theta)\}\rangle$ where $s_G(\Theta)$ is a fresh goal state, $\text{give\_up}(\Theta)$ and $\text{success}(\Theta)$ are fresh labels, $c'(\text{give\_up}(\Theta)) := 1$ and $c'(\ell) := 0$ for $\ell \neq \text{give\_up}(\Theta)$ and the transitions are defined as $T' := T \cup \{\langle s, \text{give\_up}(\Theta), s_G(\Theta)\rangle \mid$

$s \in S\} \cup \{\langle s, \text{success}(\Theta), s_G(\Theta)\rangle \mid s \in G\}$. The name of this construction represents the fact that $1 - J^*_{\text{MP}(\Theta)}(s) \in [0, 1]$ is the maximal goal probability of the state $s \in S$, i. e., the probability of terminating in the goal with a policy that maximizes this metric. In particular, $J^*_\Theta(s) \in \mathbb{R}_{\geq 0}$ if and only if $J^*_{\text{MP}(\Theta)}(s) = 0$, since an s-proper policy terminates in the goal with probability one.

For a transformation $\tau = \langle\Theta, \Theta', \sigma, \lambda\rangle$, we now investigate the corresponding extended transformation $\text{MP}(\tau) := \langle\text{MP}(\Theta), \text{MP}(\Theta'), \sigma', \lambda'\rangle$ where $\sigma'(s) := \sigma(s)$ for $s \in S$, $\sigma'(s_G(\Theta)) := s_G(\Theta')$, and $\lambda'(\ell) := \lambda(\ell)$ for $\ell \in L$, $\lambda'(\text{give\_up}(\Theta)) := \text{give\_up}(\Theta')$ and $\lambda'(\text{success}(\Theta)) := \text{success}(\Theta')$. The following is easy to see.

Theorem 2B *Let τ be a transformation. If τ satisfies any of the properties of Definition 2, then* $\text{MP}(\tau)$ *also satisfies this property.*

With the help of the MaxProb compilation and the fact that $J^*_\Theta(s) \in \mathbb{R}_{\geq 0}$ if and only if $J^*_{\text{MP}(\Theta)}(s) = 0$, we can now prove the remaining part of Theorem 2.

Theorem 2C *Let $\tau = \langle\Theta, \Theta', \sigma, \lambda\rangle$ be a transformation.*

(i) If τ is conservative, h^τ_Θ is safe and admissible.
(ii) If τ is refinable, $h^\tau_{\Theta'}$ is safe and admissible.

Proof. For (i), we can use Theorem 2B to apply Theorem 2A (i) to $\text{MP}(\tau)$ and obtain that the heuristic $h^{\text{MP}(\tau)}_{\text{MP}(\Theta)}$ is goal-aware and consistent. Note that the MaxProb compilation does not contain any dead-ends, so $h^{\text{MP}(\tau)}_{\text{MP}(\Theta)}$ is trivially safe and thus admissible. Concludingly, for all states $s \in S$, $h^{\text{MP}(\tau)}_{\text{MP}(\Theta)}(s) = J^*_{\text{MP}(\Theta')}(\sigma(s)) \leq J^*_{\text{MP}(\Theta)}(s)$. Ultimately, $J^*_{\text{MP}(\Theta)}(s) = 0$ implies $J^*_{\text{MP}(\Theta')}(\sigma(s)) = 0$, which means that $J^*_\Theta(s) \in \mathbb{R}_{\geq 0}$ implies $h^\tau_\Theta(s) = J^*_{\Theta'}(\sigma(s)) \in \mathbb{R}_{\geq 0}$. This shows that h^τ_Θ is also safe and therefore admissible.

For (ii), we likewise use Theorem 2b to apply Theorem 2a (ii) to $\text{MP}(\tau)$ and obtain that the heuristic $h^\tau_{\Theta'}$ for $\text{MP}(\Theta')$ is goal-aware and consistent. $h^{\text{MP}(\tau)}_{\text{MP}(\Theta')}$ is also trivially safe and thus admissible. This means that $h^{\text{MP}(\tau)}_{\text{MP}(\Theta')}(s') \leq J^*_{\text{MP}(\Theta')}(s')$ and therefore $J^*_{\text{MP}(\Theta)}(s) \leq J^*_{\text{MP}(\Theta')}(s')$ for all $s \in \sigma^{-1}(s')$. Ultimately, $J^*_{\text{MP}(\Theta')}(s') = 0$ implies $J^*_{\text{MP}(\Theta)}(s) = 0$ for all $s \in \sigma^{-1}(s')$, which means that $J^*_{\Theta'}(s') \in \mathbb{R}_{\geq 0}$ implies $J^*_\Theta(s) \in \mathbb{R}_{\geq 0}$ and for all states $s \in \sigma^{-1}(s')$ and thus $h^\tau_{\Theta'}(s') \in \mathbb{R}_{\geq 0}$. This shows that $h^\tau_{\Theta'}$ is also safe and admissible for refinable τ. □

Proof of Theorem 3

Theorem 3 provided a characterization of exact transformations in terms of PTS bisimulation, but the proof was only sketched. We now provide a rigorous proof of this statement.

Theorem 3 *Let Θ be a PTS and let $\tau = \langle\Theta, \Theta^{\sigma,\text{id}}, \sigma, \text{id}\rangle$. Then τ is exact if and only if $\sim_\sigma$ is a PTS bisimulation on Θ.*

Proof. Before we prove both implications, acknowledge that $s_1 \sim_\sigma s_2$ is equivalent to $\sigma(s_1) = \sigma(s_2)$ and $\delta_1 \sim_\sigma \delta_2$ is equivalent to $\sigma_\mathcal{D}(\delta_1) = \sigma_\mathcal{D}(\delta_2)$ by definition of $\sim_\sigma$.

"⇐" We show the following two properties:

(i) If $\sim_\sigma$ satisfies $\mathbf{BISIM_1}$ then τ satisfies $\mathbf{REF_G}$.
(ii) If $\sim_\sigma$ satisfies $\mathbf{BISIM_2}$ then τ satisfies $\mathbf{REF_T}$.

$\mathbf{REF_C}$ trivializes because the label mapping is the identity function and the cost functions of Θ and $\Theta^{\sigma,\text{id}}$ coincide.

$\mathbf{BISIM_1} \Rightarrow \mathbf{REF_G}$**:** Consider a goal state $s' \in \Theta^{\sigma,\text{id}}$ and let $s \in \sigma^{-1}(s')$. Because τ satisfies $\mathbf{IND_G}$, there is a goal state $\tilde{s} \in \sigma^{-1}(s')$ of Θ. In particular, $\sigma(s) = \sigma(\tilde{s}) = s'$ and we have $s \sim_\sigma \tilde{s}$. Ultimately, because $\sim_\sigma$ satisfies $\mathbf{BISIM_1}$ and $\tilde{s}$ is a goal state, s is also a goal state of Θ.

$\mathbf{BISIM_2} \Rightarrow \mathbf{REF_T}$**:** Consider a transition $\langle s', \ell, \delta'\rangle \in \Theta^{\sigma,\text{id}}$ and a state $s \in \sigma^{-1}(s')$. We must show that there exists a transition $\langle s, \ell, \delta\rangle \in \Theta$ with $\delta \in \sigma_\mathcal{D}^{-1}(\delta')$. Because τ satisfies $\mathbf{IND_T}$, there exists a transition $\langle\tilde{s}, \ell, \tilde{\delta}\rangle \in \Theta$ such that $\tilde{s} \in \sigma^{-1}(s')$ and $\tilde{\delta} \in \sigma_\mathcal{D}^{-1}(\delta')$. In particular, $s \sim_\sigma \tilde{s}$. Due to $\mathbf{BISIM_2}$, there is a transition $\langle s, \ell, \delta\rangle \in \Theta$ with $\delta \sim_\sigma \tilde{\delta}$. Ultimately, $\sigma_\mathcal{D}(\delta) = \sigma_\mathcal{D}(\tilde{\delta}) = \delta'$ and therefore $\delta \in \sigma_\mathcal{D}^{-1}(\delta')$.

"⇒" We show the following two properties:

(i) If τ satisfies $\mathbf{REF_G}$, then $\sim_\sigma$ satisfies $\mathbf{BISIM_1}$.
(ii) If τ satisfies $\mathbf{REF_T}$, then $\sim_\sigma$ satisfies $\mathbf{BISIM_2}$.

Now, let $s_1, s_2 \in \Theta$ be two states with $\sigma(s_1) = \sigma(s_2)$.

$\mathbf{REF_G} \Rightarrow \mathbf{BISIM_1}$**:** Assume that s_1 is a goal state of Θ. Because of $\mathbf{CONS_G}$, we conclude that $\sigma(s_1)$ is a goal state of $\Theta^{\sigma,\text{id}}$. From $\sigma(s_1) = \sigma(s_2)$, we have $s_2 \in \sigma^{-1}(\sigma(s_1))$. Combined with $\mathbf{REF_G}$, it follows that s_2 is a goal state of Θ.

$\mathbf{REF_T} \Rightarrow \mathbf{BISIM_2}$: Let $\langle s_1, \ell, \delta_1\rangle \in \Theta$. We must show that there exists $\langle s_2, \ell, \delta_2\rangle \in \Theta$ with $\sigma_{\mathcal{D}}(\delta_1) = \sigma_{\mathcal{D}}(\delta_2)$. Because of $\mathbf{CONS_{S+L+T}}$, we have $\langle \sigma(s_1), \ell, \sigma_{\mathcal{D}}(\delta_1)\rangle \in \Theta^{\sigma,\mathrm{id}}$. Because of $\sigma(s_1) = \sigma(s_2)$, we once again conclude $s_2 \in \sigma^{-1}(\sigma(s_1))$. Combined with $\mathbf{REF_T}$, we derive that there is a successor distribution δ_2 with $\delta_2 \in \sigma_{\mathcal{D}}^{-1}(\sigma_{\mathcal{D}}(\delta_1))$ such that $\langle s_2, \ell, \delta_2\rangle \in \Theta$. In particular, $\sigma_{\mathcal{D}}(\delta_1) = \sigma_{\mathcal{D}}(\delta_2)$. □

Proof of Theorem 5

Theorem 5 claimed that every shrink transformation is both induced and conservative. Here, we provide a detailed analysis of this statement. To this end, the following intermediate lemma will prove helpful as a simplification rule.

Lemma 1 *Let $F = (A_i)_{i\in I}$ be a factored APTS and let $\Sigma = (\mathrm{Atom}(i, \sigma_i))_{i\in I}$ be an F2FM only consisting of atomic FMs. Furthermore, let $(\alpha_i)_{i\in I}$ with $\alpha_i : E \to S_i$ be a tuple of successor mappings over the state spaces S_i of A_i and let ℓ be a label of F. It holds that $[\![\Sigma]\!]_{\mathcal{D}}(\delta_\ell(\bigotimes_{i\in I} \alpha_i)) = \delta_\ell(\bigotimes_{i\in I} \sigma_i \circ \alpha_i)$.*

Proof. Let $s' \in S'$. Firstly, since Σ consists only of atomic FMs, we have $[\![\Sigma]\!]^{-1}(s') = \times_{i\in I} \sigma_i^{-1}(s'_i)$. Secondly, for the pre-image of a product successor mapping, note that $(\bigotimes_{i\in I} \alpha_i)^{-1}(s) = \bigcap_{i\in I} \alpha_i^{-1}(s_i)$ (for arbitrary α_i). To prove the equation, we unfold the definitions of $[\![\Sigma]\!]_{\mathcal{D}}$ and δ_ℓ and make use of these two facts:

$$\begin{aligned}[\![\Sigma]\!]_{\mathcal{D}}(\delta_\ell(\bigotimes_{i\in I}\alpha_i))(s') &= \sum_{\substack{s\in\times_{i\in I}\sigma_i^{-1}(s'_i)\\ e\in\bigcap_{i\in I}\alpha_i^{-1}(s_i)}} D_\ell(e) \qquad = \sum_{e\in\bigcap_{i\in I}\alpha_i^{-1}(\sigma_i^{-1}(s'_i))} D_\ell(e)\\ &= \sum_{e\in\bigcap_{i\in I}(\sigma_i\circ\alpha_i)^{-1}(s'_i)} D_\ell(e) = \delta_\ell(\bigotimes_{i\in I}\sigma_i\circ\alpha_i)\end{aligned}$$

□

We now proceed to prove Theorem 5 in detail.

Theorem 5 *Shrink transformations are ind. abstractions.*

Proof. The properties $\mathbf{CONS_{S+L+C+G}}$ and $\mathbf{IND_{S+L+C+G}}$ are proven exactly as in the classical theory. We therefore only focus on $\mathbf{CONS_T}$ and $\mathbf{IND_T}$. In the following, let $F = (A_i)_{i\in I}$ and let $\tau = \langle F, F', \Sigma, \mathrm{id}\rangle$ be any transformation such that $\Sigma = (\mathrm{Atom}(i, \sigma_i))_{i\in I}$ is an F2FM only consisting of atomic FMs and $F' = (A'_i)_{i\in I}$ with $A'_i = A_i^{\sigma_i}$. We show that F satisfies $\mathbf{IND_T}$ and $\mathbf{CONS_T}$. Note that shrinking transformations are a special case of such transformations where only a single σ_i is not the identity function.

$\mathbf{CONS_T}$ Let $\langle s, \ell, \delta\rangle \in \Theta(\bigotimes F)$. We need to show that $\mathrm{ind}_\tau(\langle s, \ell, \delta\rangle) \in \Theta(\bigotimes F')$. By definition of $\bigotimes F$ and $\Theta(\bigotimes F)$, there are transitions $\langle s_i, \ell, \alpha_i\rangle \in A_i$ for all $i \in I$ such that $\delta = \delta_\ell(\bigotimes_{i\in I} \alpha_i)$. We have $\langle \sigma_i(s_i), \ell, \sigma_i \circ \alpha_i\rangle \in A'_i = A_i^{\sigma_i}$ by definition of $A_i^{\sigma_i}$. By definition of $\bigotimes F'$ and $\Theta(\bigotimes F')$, we have $\langle (\sigma_i(s_i))_{i\in I}, \ell, \delta_\ell(\bigotimes_{i\in I} \sigma_i \circ \alpha_i)\rangle \in \Theta(\bigotimes F')$. Ultimately, we have $[\![\Sigma]\!](s) = (\sigma_i(s_i))_{i\in I}$ by definition of $[\![\Sigma]\!]$ and $\delta_\ell(\bigotimes_{i\in I} \sigma_i \circ \alpha_i) = [\![\Sigma]\!]_{\mathcal{D}}(\delta_\ell(\bigotimes_{i\in I} \alpha_i))$ by application of Lemma 1, which shows the claim by definition of ind_τ.

$\mathbf{IND_T}$ Let $\langle s', \ell, \delta'\rangle \in \Theta(\bigotimes F')$. We need to construct a transition $\langle s, \ell, \delta\rangle \in \Theta(\bigotimes F)$ with $\mathrm{ind}_\tau(\langle s, \ell, \delta\rangle) = \langle s', \ell, \delta'\rangle$. By definition of $\bigotimes F'$ and $\Theta(\bigotimes F')$, there are transitions $\langle s'_i, \ell, \alpha'_i\rangle \in A'_i$ for all $i \in I$ such that $\delta' = \delta_\ell(\bigotimes_{i\in I} \alpha'_i)$. By definition of $A_i^{\sigma_i} = A'_i$ there exist transitions $\langle s_i, \ell, \alpha_i\rangle \in A_i$ for all $i \in I$ such that $\sigma_i(s_i) = s'_i$ and $\sigma_i \circ \alpha_i = \alpha'_i$. By definition of $\bigotimes F$ and $\Theta(\bigotimes F)$, we conclude $\langle (s_i)_{i\in I}, \ell, \delta_\ell(\bigotimes_{i\in I} \alpha_i)\rangle \in \Theta(\bigotimes F)$. Last but not least, note that $[\![\Sigma]\!]((s_i)_{i\in I}) = (\sigma_i(s_i))_{i\in I} = s'$ and $[\![\Sigma]\!]_{\mathcal{D}}(\delta_\ell(\bigotimes_{i\in I} \alpha_i)) = \delta_\ell(\bigotimes_{i\in I} \alpha'_i)$ by Lemma 1, which shows the claim by definition of ind_τ. □

Proof of Theorem 6

In Theorem 6, we made the statement that a shrink transformation that uses a local abstraction generated from an APTS bisimulation is an exact transformation, which strengthens Theorem 5 for such shrinking strategies. In the following, we give a complete and formal proof for this statement.

Theorem 6 *APTS bisimulation-based shrinking is exact.*

Proof. Let $F = (A_i)_{i\in I}$ with $A_i = \langle S_i, L, c, E, D, T_i, G_i\rangle$ and $\bigotimes F = \langle S, L, c, E, D, T, G\rangle$. Let $\sim$ be an APTS bisimulation on $A_k \in F$ and let $\tau = \langle F, F', \Sigma, \mathrm{id}\rangle$ be the shrinking transformation for $\sigma_\sim$ and A_k.

By Theorem 5, every shrinking transformation is induced conservative. In particular, $\Theta(\bigotimes F') = \Theta(\bigotimes F)^{[\![\Sigma]\!],\mathrm{id}}$. By applying Theorem 3, we can therefore conclude that τ is exact if and only if the induced equivalence relation $\sim_{[\![\Sigma]\!]}$ of $[\![\Sigma]\!]$ is a PTS bisimulation on $\Theta(\bigotimes F)$.

To show this, let $s, t \in \Theta(\bigotimes F)$ such that $s \sim_{[\![\Sigma]\!]} t$, i. e., $[\![\Sigma]\!](s) = [\![\Sigma]\!](t)$ in the following. Note that this implies $s_i = t_i$ for $i \neq k$ and $s_k \sim t_k$ by definition of Σ and $\sigma_\sim$.

BISIM$_1$ Assume $s \in G$. We must show that $t \in G$. By definition of $\bigotimes F$, we get $s_i \in G_i$ for all $i \in I$. For $i \neq k$, we have $t_i = s_i \in G_i$. Because $\sim$ is a PTS bisimulation on A_k and since $s_k \in G_k$ and $s_k \sim t_k$, we also have $t_k \in G_k$ by **BISIM$_1$** for $\sim$. All in all, $t \in G$ by definition of $\bigotimes F$.

BISIM$_2$ Let $\langle s, \ell, \delta \rangle \in \Theta(\bigotimes F)$. By definition of $\bigotimes F$ and $\Theta(\bigotimes F)$, there exist transitions $\langle s_i, \ell, \alpha_i \rangle \in A_i$ with $\delta = \delta_\ell(\bigotimes_{i \in I} \alpha_i)$. Since $\sim$ is an APTS bisimulation on A_k and $s_k \sim t_k$, there exists a transition $\langle t_k, \ell, \beta_k \rangle \in A_k$ such that $\alpha_k \sim \beta_k$ by **BISIM$_2$** for $\sim$. Now define $\beta_i := \alpha_i$ for $i \neq k$. By definition of $\bigotimes F$ and $\Theta(\bigotimes F)$, we have $\langle t, \ell, \delta_\ell(\bigotimes_{i \in I} \beta_i) \rangle \in \Theta(\bigotimes F)$.

It remains to show that $\delta_\ell(\bigotimes_{i \in I} \alpha_i) \sim_{[\![\Sigma]\!]} \delta_\ell(\bigotimes_{i \in I} \beta_i)$, i. e., $\sum_{u \in C} \delta_\ell(\bigotimes_{i \in I} \alpha_i)(u) = \sum_{u \in C} \delta_\ell(\bigotimes_{i \in I} \beta_i)(u)$ for all equivalence classes $C \in S/{\sim_{[\![\Sigma]\!]}}$. Let $C \in S/{\sim_{[\![\Sigma]\!]}}$ be an equivalence class. We start by applying the definition of δ_ℓ:

$$\sum_{u \in C} \delta_\ell(\bigotimes_{i \in I} \beta_i)(u) = \sum_{u \in C} \sum_{e \in \bigcap_{i \in I} \beta_i^{-1}(u_i)} D_\ell(e).$$

We can further partition the innermost sum by first considering the equivalence classes of $\simeq_{A_k,\ell}$, which gives us

$$\ldots = \sum_{u \in C} \sum_{\substack{D \in E/\simeq_{A_k,\ell} \\ e \in D \cap \bigcap_{i \in I} \beta_i^{-1}(u_i)}} D_\ell(e).$$

Now, note that every equivalence class $D \in E/{\simeq_{A_k,\ell}}$ is either completely contained in the intersection $\bigcap_{i \in I \setminus \{k\}} \beta_i^{-1}(u_i)$ or completely disjoint from it by definition of $\simeq_{A_k,\ell}$. This allows us to restrict to those equivalence classes which are contained in the intersection.

$$\ldots = \sum_{u \in C} \sum_{\substack{D \in E/\simeq_{A_k,\ell} \\ D \subseteq \bigcap_{i \in I \setminus \{k\}} \beta_i^{-1}(u_i)}} \sum_{e \in D \cap \beta_k^{-1}(u_k)} D_\ell(e)$$

By definition of Σ and $\sim_F$, we can write the equivalence class C as $C = \bigtimes_{i \in I} C_i$, where $C_k \in S_k/{\sim}$ and $C_i = \{\tilde{u}_i\}$ for fixed states $\tilde{u}_i \in A_i$, $i \neq k$. The outermost sum therefore only needs to consider the k-th component and can be moved inwards.

$$\ldots = \sum_{\substack{D \in E/\simeq_{A_k,\ell} \\ D \subseteq \bigcap_{i \in I \setminus \{k\}} \beta_i^{-1}(\tilde{u}_i)}} \sum_{u_k \in C_k} \sum_{e \in D \cap \beta_k^{-1}(u_k)} D_\ell(e)$$

Next, we apply the facts $\alpha_k \sim \beta_k$ and $\alpha_i = \beta_i$ for $i \neq k$.

$$\ldots = \sum_{\substack{D \in E/\simeq_{A_k,\ell} \\ D \subseteq \bigcap_{i \in I \setminus \{k\}} \alpha_i^{-1}(\tilde{u}_i)}} \sum_{u_k \in C_k} \sum_{e \in D \cap \alpha_k^{-1}(u_k)} D_\ell(e)$$

Lastly, we apply the previous transformations in reverse.

$$\ldots = \sum_{u \in C} \sum_{\substack{D \in E/\simeq_{A_k,\ell} \\ D \subseteq \bigcap_{i \in I \setminus \{k\}} \alpha_i^{-1}(u_i)}} \sum_{e \in D \cap \alpha_k^{-1}(u_k)} D_\ell(e) = \sum_{u \in C} \sum_{\substack{D \in E/\simeq_{A_k,\ell} \\ e \in D \cap \bigcap_{i \in I} \alpha_i^{-1}(u_i)}} D_\ell(e)$$

$$= \sum_{u \in C} \sum_{e \in \bigcap_{i \in I} \alpha_i^{-1}(u_i)} D_\ell(e) = \sum_{u \in C} \delta_\ell(\bigotimes_{i \in I} \alpha_i)(u)$$

□

Proof of Theorem 7

In Theorem 7, we made the claim that all merge transformations are isomorphisms, and therefore only cause a renaming of the states of the represented state space.

Theorem 7 *All merge transformations are isomorphisms.*

Proof. Let $\tau = \langle F, F', \Sigma, \mathrm{id}\rangle$ be a merge transformation of $F = (A_i)_{i\in I}$ into $F' = (A'_i)_{i\in I'}$ for $j, k \in I$ with $j \neq k$, where $I' = I \setminus \{j,k\} \cup \{\langle j,k\rangle\}$. We need to show that τ is an isomorphism. Since merge transformations only affect the (goal) states and transitions of $\bigotimes F$, it suffices to show that $[\![\Sigma]\!]$ is bijective, that $[\![\Sigma]\!](G) = G'$ for the goal states G and G' of $\bigotimes F$ and $\bigotimes F'$, and that $T' = \{\langle [\![\Sigma]\!](s), \ell, [\![\Sigma]\!] \circ \alpha\rangle \mid \langle s, \ell, \alpha\rangle \in T\}$ for the transitions T and T' of $\bigotimes F$ and $\bigotimes F'$. Since Σ, G and G' are exactly defined as in the classical theory of merge transformations, the former two claims are already proven in the classical case. We only need to focus on the transitions. We show both inclusions seperately.

"$\supseteq$" Let $\langle s, \ell, \beta\rangle \in \bigotimes F$. We need to show $\langle [\![\Sigma]\!](s), \ell, [\![\Sigma]\!] \circ \beta\rangle \in \bigotimes F'$. By definition of $\bigotimes F$, there exist transitions $\langle s_i, \ell, \alpha_i\rangle \in A_i$ for each $i \in I$ such that $\beta = \bigotimes_{i\in I} \alpha_i$. Furthermore, since $\langle s_j, \ell, \alpha_j\rangle \in A_j$ and $\langle s_k, \ell, \alpha_k\rangle \in A_k$, we conclude that $\langle\langle s_j, s_k\rangle, \ell, \alpha_j \otimes \alpha_k\rangle \in A_j \otimes A_k$ by definition of $A_j \otimes A_k$. We now define

$$s' := (s_i)_{i\in I\setminus\{j,k\}} \sqcup \{\langle j,k\rangle \mapsto \langle s_j, s_k\rangle\} \qquad \alpha' := (\alpha_i)_{i\in I\setminus\{j,k\}} \sqcup \{\langle j,k\rangle \mapsto \alpha_j \otimes \alpha_k\}.$$

We have that $\langle s'_i, \ell, \alpha'_i\rangle \in A'_i$ for every $i \in I'$ by definition, and therefore $\langle s', \ell, \bigotimes_{i\in I'} \alpha'_i\rangle \in \bigotimes F'$ by definition of $\bigotimes F'$. We clearly have $[\![\Sigma]\!](s) = s'$ by definition of Σ and s'. It is left to show that $[\![\Sigma]\!] \circ \bigotimes_{i\in I} \alpha_i = \bigotimes_{i\in I'} \alpha'_i$. To this end, let e be an label effect. We have:

$$\begin{aligned}([\![\Sigma]\!] \circ \bigotimes_{i\in I} \alpha_i)(e) = [\![\Sigma]\!]((\alpha_i(e))_{i\in I}) &= (\alpha_i(e))_{i\in I\setminus\{j,k\}} \sqcup \{\langle j,k\rangle \mapsto \langle \alpha_j(e), \alpha_k(e)\rangle\}\\ &= (\alpha_i(e))_{i\in I\setminus\{j,k\}} \sqcup \{\langle j,k\rangle \mapsto (\alpha_j \otimes \alpha_k)(e)\}\\ &= (\alpha'_i(e))_{i\in I\setminus\{j,k\}} \sqcup \{\langle j,k\rangle \mapsto (\alpha'_{\langle j,k\rangle})(e)\}\\ &= (\alpha'_i(e))_{i\in I'} = (\bigotimes_{i\in I'} \alpha'_i)(e).\end{aligned}$$

"$\subseteq$" Let $\langle s', \ell, \beta'\rangle \in \bigotimes F'$. By definition of $\bigotimes F'$, there exist transitions $\langle s'_i, \ell, \alpha'_i\rangle \in A'_i$ for each $i \in I'$ such that $\beta' = \bigotimes_{i\in I'} \alpha'_i$. From $\langle s'_{\langle j,k\rangle}, \ell, \alpha'_{\langle j,k\rangle}\rangle \in A'_{\langle j,k\rangle} = A_j \otimes A_k$ and the definition of $A_j \otimes A_k$, there exist $\langle \tilde{s}_j, \ell, \tilde{\alpha}_j\rangle \in A_j$ and $\langle \tilde{s}_k, \ell, \tilde{\alpha}_k\rangle \in A_k$ such that $s'_{\langle j,k\rangle} = \langle \tilde{s}_j, \tilde{s}_k\rangle$ and $\alpha'_{\langle j,k\rangle} = \tilde{\alpha}_j \otimes \tilde{\alpha}_k$. We now define

$$s := (s'_i)_{i\in I\setminus\{j,k\}} \sqcup \{j \mapsto \tilde{s}_j\} \sqcup \{k \mapsto \tilde{s}_k\} \qquad \alpha := (\alpha'_i)_{i\in I\setminus\{j,k\}} \sqcup \{j \mapsto \tilde{\alpha}_j\} \sqcup \{k \mapsto \tilde{\alpha}_k\}.$$

By definition, we have $\langle s_i, \ell, \alpha_i\rangle \in A_i$ for all $i \in I$ and therefore $\langle s, \ell, \bigotimes_{i\in I} \alpha_i\rangle \in \bigotimes F$ by definition of $\bigotimes F$. It is clear that $[\![\Sigma]\!](s) = s'$ by definition of Σ and s. It is left to show that $[\![\Sigma]\!] \circ \bigotimes_{i\in I} \alpha_i = \bigotimes_{i\in I'} \alpha'_i$. To this end, we have:

$$\begin{aligned}([\![\Sigma]\!] \circ \bigotimes_{i\in I} \alpha_i)(e) = [\![\Sigma]\!]((\alpha_i(e))_{i\in I}) &= (\alpha_i(e))_{i\in I\setminus\{j,k\}} \sqcup \{\langle j,k\rangle \mapsto \langle \tilde{\alpha}_j(e), \tilde{\alpha}_k(e)\rangle\}\\ &= (\alpha_i(e))_{i\in I\setminus\{j,k\}} \sqcup \{\langle j,k\rangle \mapsto (\tilde{\alpha}_j \otimes \tilde{\alpha}_k)(e)\}\\ &= (\alpha'_i(e))_{i\in I\setminus\{j,k\}} \sqcup \{\langle j,k\rangle \mapsto (\alpha'_{\langle j,k\rangle})(e)\}\\ &= (\alpha'_i(e))_{i\in I'} = (\bigotimes_{i\in I'} \alpha'_i)(e).\end{aligned}$$

□

Proof of Theorem 8

In Theorem 8, we investigated the basic properties of label reductions. The proof given in the paper contained a final claim that we deal with here. For the sake of simplicity, we including here the steps already mentioned in the paper.

Theorem 8 *Label reductions are abstractions and satisfy* $\mathbf{IND_{S+L+C+G}}$ *and* $\mathbf{REF_G}$*. Furthermore, they satisfy* $\mathbf{REF_C}$ *if and only if only labels with the same costs are reduced and satisfy* $\mathbf{IND_T}$ *if and only if they satisfy* $\mathbf{REF_T}$*.*

Proof. The proofs for properties $\mathbf{CONS_{S+L+C+G}}$, $\mathbf{IND_{S+L+C+G}}$ and $\mathbf{REF_G}$, and the necessary and sufficient condition of $\mathbf{REF_C}$ are given in the classical theory. Since $[\![\Sigma]\!] = \mathrm{id}$ we have $\bigcap_{s\in[\![\Sigma]\!]^{-1}(s')} \mathrm{ind}_\tau(T(s)) = \mathrm{ind}_\tau(s')$ and so $\mathbf{IND_T}$ and $\mathbf{REF_T}$ collapse to a common statement.

For $\mathbf{CONS_T}$, let $\langle s, \ell, \delta\rangle \in \Theta(\bigotimes F)$. By definition of $\Theta(A)$ and the product, there is some α such that $\delta = \delta_\ell(\bigotimes_{i\in I} \alpha_i)$ and $\langle s_i, \ell, \alpha_i\rangle \in A_i$ for all $i \in I$. By definition of $A_i^{\lambda,\epsilon}$, we get $\langle s_i, \lambda(\ell), \alpha_i \circ \epsilon_\ell^{-1}\rangle \in A_i^{\lambda,\epsilon}$ for all $i \in I$. By definition of the product and $\Theta(A)$, $\langle s, \lambda(\ell), \delta_{\lambda(\ell)}(\bigotimes_{i\in I} \alpha_i \circ \epsilon_\ell^{-1})\rangle \in \Theta(\bigotimes F')$. It is left to show that $\delta_\ell(\bigotimes_{i\in I} \alpha_i) = \delta_{\lambda(\ell)}(\bigotimes_{i\in I} \alpha_i \circ \epsilon_\ell^{-1})$.

To this end, let $s \in \bigotimes F$. We have

$$\delta_\ell(\bigotimes_{i\in I} \alpha_i)(s) = \sum_{e \in \bigcap_{i\in I} \alpha_i^{-1}(s)} D_\ell(e) = \sum_{e \in \bigcap_{i\in I} \alpha_i^{-1}(s)} (D'_\ell \circ \epsilon_\ell)(e)$$

by definition of δ_ℓ and D'. We now move ϵ_ℓ to the range of the sum, so we get:

$$\ldots \quad = \sum_{e' \in \epsilon_\ell(\bigcap_{i\in I} \alpha_i^{-1}(s))} D'_\ell(e')$$

Note that ϵ_ℓ is bijective. For injective functions, the function image commutes with intersections, so we obtain:

$$\ldots \quad = \sum_{e' \in \bigcap_{i\in I} \epsilon_\ell(\alpha_i^{-1}(s))} D'_\ell(e')$$

Lastly, we apply the inverse on a function composition and finally use the definition of δ_ℓ to conclude the claim.

$$\ldots \quad = \sum_{e' \in \bigcap_{i\in I} (\alpha_i \circ \epsilon_\ell^{-1})^{-1}(s)} D'_\ell(e') = \delta_{\lambda(\ell)}(\bigotimes_{i\in I} \alpha_i \circ \epsilon_\ell^{-1})(s).$$

□

Proof of Theorem 9

Theorem 9 covered exact label reduction.

Theorem 9 *A label reduction of a FAPTS F that combines exactly two labels ℓ_1, ℓ_2 of F is exact if $c(\ell_1) = c(\ell_2)$ and either*

(A) ℓ_1 ϵ-subsumes ℓ_2 in all $A \in F$ or vice versa,
(B) ℓ_1 and ℓ_2 are (A, ϵ)-combinable for some $A \in F$ or
(C) ℓ_1 and ℓ_2 are dead for some $A \in F$.

We split the proof of Theorem 9 into two subtheorems, Theorem 9A and Theorem 9B, from which Theorem 9 follows as a corollary.

Theorem 9A *A label reduction for $\lambda : L \to L'$ and ϵ of an FAPTS $F = (A_i)_{i\in I}$ that combines exactly two labels $\ell_1, \ell_2 \in L$ satisfies*

$$\bigtimes_{i\in I} \bigcup_{\ell \in \lambda^{-1}(\ell')} T_\epsilon(A_i, \ell) \subseteq \bigcup_{\ell \in \lambda^{-1}(\ell')} \bigtimes_{i\in I} T_\epsilon(A_i, \ell)$$

for all $\ell' \in L'$ if one of the following conditions holds:

(A) ℓ_1 ϵ-subsumes ℓ_2 in all $A \in F$ or vice versa,
(B) ℓ_1 and ℓ_2 are (A, ϵ)-combinable for some $A \in F$ or
(C) ℓ_1 and ℓ_2 are dead for some $A \in F$.

Proof. The inclusion is trivial for $\ell' \neq \ell_{12}$. For ℓ_{12}, the inclusion collapses to:

$$\bigtimes_{i\in I} \Big(T_\epsilon(A_i, \ell_1) \cup T_\epsilon(A_i, \ell_2)\Big) \subseteq \bigtimes_{i\in I} T_\epsilon(A_i, \ell_1) \cup \bigtimes_{i\in I} T_\epsilon(A_i, \ell_2).$$

We now prove this inclusion for every case.

Case (A) Assume ℓ_1 ϵ-subsumes ℓ_2 in all factors, the other case is symmetric. We have

$$\bigtimes_{i\in I} \Big(T_\epsilon(A_i, \ell_1) \cup T_\epsilon(A_i, \ell_2)\Big) = \bigtimes_{i\in I} T_\epsilon(A_i, \ell_1) \subseteq \bigtimes_{i\in I} T_\epsilon(A_i, \ell_1) \cup \bigtimes_{i\in I} T_\epsilon(A_i, \ell_2).$$

Case (B) Choose $k \in I$ such that ℓ_1 and ℓ_2 are (A_k, ϵ)-combinable. Then:

$$\begin{aligned}
\bigtimes_{i\in I} \Big(T_\epsilon(A_i, \ell_1) \cup T_\epsilon(A_i, \ell_2)\Big) &= \bigtimes_{i\in I} \begin{Bmatrix} T_\epsilon(A_i, \ell_1) & \text{if } i = k \\ T_\epsilon(A_i, \ell_1) \cup T_\epsilon(A_i, \ell_2) & \text{if } i \neq k \end{Bmatrix} \\
&= \bigtimes_{i\in I} T_\epsilon(A_i, \ell_1) \cup \bigtimes_{i\in I} \begin{Bmatrix} T_\epsilon(A_i, \ell_1) & \text{if } i = k \\ T_\epsilon(A_i, \ell_2) & \text{if } i \neq k \end{Bmatrix} \\
&= \bigtimes_{i\in I} T_\epsilon(A_i, \ell_1) \cup \bigtimes_{i\in I} T_\epsilon(A_i, \ell_2).
\end{aligned}$$

Case (C) Let $k \in I$ such that ℓ_1, ℓ_2 are dead in $A_k \in F$. Then $T_\epsilon(A_k, \ell_1) = \emptyset = T_\epsilon(A_i, \ell_2)$ and both sides are empty. $\square$

Theorem 9B *A label reduction for $\lambda : L \to L'$ and ϵ of an FAPTS $F = (A_i)_{i \in I}$ with labels L that combines exactly two labels $\ell_1, \ell_2 \in L$ satisfies* $\mathbf{IND_T}$ *if for all $\ell' \in L'$:*

$$\bigtimes_{i \in I} \bigcup_{\ell \in \lambda^{-1}(\ell')} T_\epsilon(A_i, \ell) \subseteq \bigcup_{\ell \in \lambda^{-1}(\ell')} \bigtimes_{i \in I} T_\epsilon(A_i, \ell)$$

Proof. We prove the statement by contraposition. Assume $\mathbf{IND_T}$ does not hold. By assumption, let $\langle s, \ell', \delta \rangle \in \Theta(\bigotimes F')$ be a transition such that there exists no transition $\langle s, \ell, \delta \rangle \in \Theta(\bigotimes F)$ with $\ell \in \lambda^{-1}(\ell')$. By definition of the product, F' and $\Theta(\bigotimes F')$, we can decompose δ into $\delta = \delta_{\ell'}(\bigotimes_{i \in I} \alpha_i \circ \epsilon_{\ell'}^{-1})$. Now, if it were the case that $\langle s, \ell, \bigotimes_{i \in I} \alpha_i \rangle \in \bigotimes F$, for some $\ell \in \lambda^{-1}(\ell')$, then we would have $\delta_\ell(\bigotimes_{i \in I} \alpha_i) = \delta_{\ell'}(\bigotimes_{i \in I} \alpha_i \circ \epsilon_{\ell'}^{-1})$. We therefore have $\langle s, \ell, \bigotimes_{i \in I} \alpha_i \rangle \notin \bigotimes F$ for all $\ell \in \lambda^{-1}(\ell')$.

Now, let $p = (\langle s_i, \alpha_i \circ \epsilon_{\ell'}^{-1} \rangle)_{i \in I}$. We will show that $p \in \mathrm{LHS} := \bigtimes_{i \in I} \bigcup_{\ell \in \lambda^{-1}(\ell')} T_\epsilon(A_i, \ell)$, but $p \notin \mathrm{RHS} := \bigcup_{\ell \in \lambda^{-1}(\ell')} \bigtimes_{i \in I} T_\epsilon(A_i, \ell)$, thereby showing that the inclusion in the assumption does not hold.

$p \in \mathrm{LHS}$ We have $\langle s, \ell', \bigotimes_{i \in I} \alpha_i \circ \epsilon_{\ell'}^{-1} \rangle \in \bigotimes F'$. From the definition of the product, this means $\langle s_i, \ell', \alpha_i \circ \epsilon_{\ell'}^{-1} \rangle \in A_i'$ for all $i \in I$. From the definition of label reduction, this means that for all $i \in I$, there exists a label $\ell_i \in \lambda^{-1}(\ell')$ such that $\langle s_i, \ell_i, \alpha_i \rangle \in A_i$, which means $\langle s_i, \alpha_i \circ \epsilon_{\ell'}^{-1} \rangle \in T_\epsilon(A_i, \ell_i)$. Hence, $\langle s_i, \alpha_i \circ \epsilon_{\ell'}^{-1} \rangle \in \bigcup_{\ell \in \lambda^{-1}(\ell')} T_\epsilon(A_i, \ell)$ for all $i \in I$. Ultimately, $p \in \bigtimes_{i \in I} \bigcup_{\ell \in \lambda^{-1}(\ell')} T_\epsilon(A_i, \ell)$, by definition of p.

$p \notin \mathrm{RHS}$ Consider a label $\ell \in \lambda^{-1}(\ell')$. We know that $\langle s, \ell, \bigotimes_{i \in I} \alpha_i \rangle \notin \bigotimes F$. By definition of the product, there must be a $k \in I$ such that $\langle s_k, \ell, \alpha_k \rangle \notin A_k$ and therefore $\langle s_k, \alpha_k \circ \epsilon_\ell^{-1} \rangle \notin T_\epsilon(A_k, \ell)$. We obtain $p \notin \bigtimes_{i \in I} T_\epsilon(A_i, \ell)$ by definition of p. Since ℓ was arbitrary, this shows $p \notin \bigcup_{\ell \in \lambda^{-1}(\ell')} \bigtimes_{i \in I} T_\epsilon(A_i, \ell)$. $\square$