

Interactive Robot Transition Repair With SMT

Jarrett Holtz, Arjun Guha, and Joydeep Biswas
 University of Massachusetts Amherst
 {jaholtz,arjun,joydeepb}@cs.umass.edu

Abstract

Complex robot behaviors are often structured as state machines, where states encapsulate actions and a transition function switches between states. Since transitions depend on physical parameters, when the environment changes, a roboticist has to painstakingly readjust the parameters to work in the new environment. We present *interactive SMT-based Robot Transition Repair* (SRTR): instead of manually adjusting parameters, we ask the roboticist to identify a few instances where the robot is in a wrong state and what the right state should be. An automated analysis of the transition function 1) identifies adjustable parameters, 2) converts the transition function into a system of logical constraints, and 3) formulates the constraints and user-supplied corrections as a MaxSMT problem that yields new parameter values. We show that SRTR finds new parameters 1) quickly, 2) with few corrections, and 3) that the parameters generalize to new scenarios. We also show that a SRTR-corrected state machine can outperform a more complex, expert-tuned state machine.

1 Introduction

Complex robot control software is typically structured as a state machine, where each state encapsulates a feedback controller. Even if each state is correct, the transitions between states depend on parameters that are hard to get right, even for experienced roboticists. It is very common for parameter values to work in simulation but fail in the real world, to work in one physical environment but fail in another, or to work on one robot but fail on another. For example, Figure 1 shows the trajectory of a robotic soccer player as it tries to kick a moving ball. A very small change to its parameter values determines whether or not it succeeds.

Even a simple robot may have a large parameter space, which makes exhaustive-search impractical. Moreover, robot performance is usually non-convex with respect to parameter values, which makes general optimization techniques susceptible to local minima. For some cases, there exist calibration procedures to adjust parameters automatically (*e.g.*, [Holtz and Biswas, 2017]), but these are not general procedures.

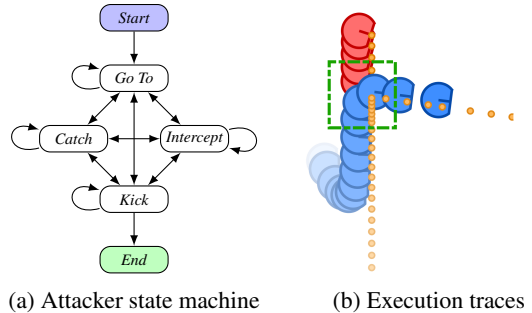


Figure 1: A robot soccer attacker a) state machine, with b) successful (blue) and unsuccessful (red) traces to *Intercept* a ball (orange) and *Kick* at the goal. The green box isolates the error: the successful trace transitions to *Kick*, the unsuccessful trace remains in *Intercept*.

Therefore, roboticists usually resort to adjusting parameters manually—a tedious task that can result in poor performance.

However, consider the following observations: a roboticist debugging a robot can frequently identify when something goes wrong, and what the robot should have done instead. When the robot control software is structured as a state machine, this corresponds to identifying when the robot is in the wrong state and what the correct state should be. This is a *partial specification* of expected behavior: *i.e.*, it does not enumerate a complete sequence of states, the parameters to adjust, how to adjust them, or even identify all errors. In this paper, we show that partial specification of this kind are adequate inputs for an automatic parameter repair procedure.

We present *Interactive SMT-based Robot Transition Repair* (SRTR). During execution, SRTR logs the execution trace of the transition function. After execution, the roboticist examines the trace and provides a handful of corrections. SRTR then converts the set of corrections and the transition function into a logical formula for a MaxSMT solver. A solution to this MaxSMT problem minimizes adjustments to the parameters, while maximizing the number of satisfied corrections.

Our experimental results demonstrate that SRTR 1) is more computationally tractable than exhaustive parameter search; 2) is scalable; 3) generalizes to unseen situations; and 4) when applied to simple robot soccer attacker, can outperform a more complicated, expert-tuned attacker that won the lower bracket finals at *RoboCup 2017*.

2 Related Work

Finite state machines (FSMs) can specify robot behavior and enable robust autonomy. RoboChart [Miyazawa *et al.*, 2017] uses FSMs to verify liveness, timeliness, and other properties of robots. The HAMQ-INT algorithm [Bai and Russell, 2017] uses reinforcement learning to discover internal transitions and adapt to unexpected scenarios. SRTR also leverages FSMs, but focuses on correcting transition errors that occur when the robot is deployed in a new environment.

Robot behavior can also be repaired by dynamic synthesis of new control structures, such as automatic synthesis of new FSMs [Wong *et al.*, 2014], or synthesis of code from a context-free motion grammar with parameters derived from human-inspired control [Dantam *et al.*, 2013]. When automated synthesis is intractable, a user-generated specification in a domain-specific language can be used to synthesize a plan [Nedunuri *et al.*, 2014]. SRTR assumes that the FSM does not need new states or transitions, but that failures are due to *incorrect triggering* of the transition function arising from incorrect environment-dependent parameter values.

Robot behaviors rely on environment-dependent parameters for robust and accurate execution. If a precise model of the dependency between parameters and behaviors is available, it may be possible to design a *calibration procedure* that executes a specific sequence of actions and to recover correct parameter values (*e.g.*, [Holtz and Biswas, 2017]). If a calibration procedure cannot be designed, but the effect of parameters is well-understood, it may be possible to optimize for the parameters using a functional model [Cano *et al.*, 2016]. Model-based diagnosis can diagnose faulty parameters [Reiter, 1987] if the behavior of the robot in its environment can be formally defined. SRTR can repair parameters even without a descriptive model of the robot’s behavior.

Human input can help overcome the limitations of autonomous algorithms [Kamar, 2016; Nashed and Biswas, 2018]. Learning from demonstration (LfD) [Argall *et al.*, 2009] and inverse reinforcement learning (IRL) [Abbeel and Ng, 2011] allow robots to learn new behaviors from human demonstrations. LfD can also overcome model errors by correcting portions of the state space [Meriçli *et al.*, 2012]. These approaches require demonstrations in the full high-dimensional state space of the robot, which can be tedious for users to provide. When human demonstration does not specify *why* an action was applied to a state, it can be hard to generalize to a new situation. SRTR generalizes to new scenarios since it infers dependencies from the code of the transition function. Furthermore, it requires only a partial specification of correct behavior.

Domain-specific languages (DSLs) allow users to specify high-level behavior using abstractions such as Instruction Graphs [Meriçli *et al.*, 2014]. Unlike DSLs that provide a complete specification of robot behavior, SRTR only requires sparse corrections that partially-specify expected behavior.

DirectFix [Mechtaev *et al.*, 2015] formulates program repair as a MaxSMT problem and deems a program fixed when all tests pass. Physical robots don’t have deterministic test cases and user-provided corrections can be contradictory. SRTR uses MaxSMT to minimize changes and max-

imize the number of satisfied corrections. There are other domains that are not amenable to unit-testing. For example, Tortoise [Weiss *et al.*, 2017] propagates system configuration fixes from the shell to a system configuration specification. However, Tortoise requires a complete fix where SRTR only requires a partial specification.

Programming by example synthesizes programs from a small number of examples [Gulwani, 2011] and can also support noisy data [Devlin *et al.*, 2017]. *Program templates* can help synthesize task and motion plans [Nedunuri *et al.*, 2014]. SRTR does not synthesize new program structure, but focuses on making minimal parameter adjustments using MaxSMT to satisfy user corrections.

3 Background

We use a real-world example to motivate SRTR: a robot soccer attacker that 1) goes to the ball if the ball is stopped, 2) intercepts the ball if it is moving away from the attacker, 3) catches the ball if it is moving toward the attacker, and 4) kicks the ball at the goal once the attacker has control of the ball. Each of these sub-behaviors is a distinct, self-contained feedback controller (*e.g.*, ball interception [Biswas *et al.*, 2014] or omnidirectional time optimal control [Kalmár-Nagy *et al.*, 2004]). At each time-step, the attacker 1) switches to a new controller if necessary and 2) invokes the current controller to produce new outputs. We represent the attacker as a *robot state machine* (RSM), where each state represents a controller (Figure 1a).

In this paper, we assume that the output of each controller is nominally correct: there may be minor *performance* degradation when environmental factors change, but we assume that they are *convergent*, and will eventually produce the correct result. However, environmental factors also affect the transition function, which transfers execution from one controller to another. For example, the friction coefficient between the ball and the carpet affects: when the attacker transitions from *Intercept* to *Kick*; and the mass of the ball affects when the attacker transitions from *Kick* to *Done*. These factors vary from one environment to another. Since transition functions do not have any self-correcting mechanisms, robots are prone to behaving incorrectly when their parameters are incorrect for the given environment.

3.1 Robot State Machines

A robot state machine (RSM) is a discrete-time Mealy machine that is extended with continuous inputs, outputs, and program variables. Formally, an RSM is a 9-tuple $\langle S, S_0, S_F, V, V_0, Y, U, T, G \rangle$, where S is the finite set of states, $S_0 \in S$ is the start state, $S_F \in S$ is the end state, $V \in \mathbb{R}^m$ is the set of program variable values, $V_0 \in \mathbb{R}^m$ are the initial values of the program variables, $Y \in \mathbb{R}^n$ are the continuous outputs, $U \in \mathbb{R}^l$ are the continuous actuation inputs, $T : S \times Y \times V \rightarrow S$ is the transition function, and $G : S \times Y \times V \rightarrow U \times V$ is the emission function. At each time step t , the RSM first uses the transition function to select a state and then the emission function to run the controller associated with that state. The transition function can only update the current state, whereas the emission function can update program variables and produce outputs.

3.2 Transition Errors

Figure 1b shows two traces of the attacker. In the blue trace, the attacker correctly intercepts the moving ball and kicks it at the goal. But, in the red trace, the attacker fails to kick: it remains stuck in the *Intercept* state and never transitions to *Kick*. Over the course of several trials (e.g., a robot soccer game), we may find that the attacker only occasionally fails to kick. When this occurs, it is usually the case that the high-level structure of the transition function is correct, but that the values of the parameters need to be adjusted. Unfortunately, since there are 11 real-valued parameters in the full attacker RSM, the search space is large.

To efficiently search for new parameter values, we need to reason about the structure of the transition function. To do so, the next section describes how we systematically convert it to a formula in propositional logic, extended with arithmetic operators. This formula encodes the structure of the transition function along with constraints from the user corrections, and allows an SMT solver to efficiently find new parameters to correct the errant transition(s).

4 Interactive Robot Transition Repair

The SRTR algorithm has four inputs: 1) the transition function, 2) a map from parameters to their values, 3) an execution trace, and 4) a set of user-provided corrections. The result of SRTR is a corrected parameter map that maximizes the number of corrections satisfied and minimizes the changes to the input parameter map. (The tradeoff between these objectives is a hyperparameter.)

SRTR has three major steps. 1) For each user-provided correction, it *partially evaluates* the transition function for the inputs and variable values at the time of correction, yielding *residual transition functions* (Section 4.2). 2) Finally, it uses the residual transition functions to formulate an optimization problem for an off-the-shelf MaxSMT solver, for this paper we use Z3 [Björner *et al.*, 2015]. The solution to this problem is an adjustment to the parameter values (Section 4.3).

To illustrate the SRTR algorithm, we present as a running example a simplified attacker RSM that is only capable of handling a stationary ball on the field (Figure 2). Therefore, the RSM has four states (Start, GoTo, Kick, and End) and its transition function (Figure 3a) has four parameters (aimMargin_p, maxDist_p, viewAng_p, and kickTimeout_p). From an execution trace of the RSM (Figure 3b), we consider an example where at time-step $t = 5$, the user identifies that the transition function produces an incorrect result: instead of GoTo, it should have returned Kick. With this example in mind, we present how SRTR uses the transition function code, an execution trace, and a correction to identify that an adjustment to just one of the parameters, maxDist_p, is sufficient to satisfy this correction (Figure 3c).

4.1 Transition Functions and SRTR Inputs

To abstract away language-specific details of our repair procedure, we present SRTR for an idealized imperative language that only has features essential for writing transition functions. Figure 4 lists the syntax for a transition function written in SRTR-repairable form using a notation that is close

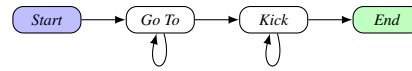


Figure 2: Simplified attacker RSM.

to standard BNF. In general, the transition function consists of sequences of statements comprised of expressions over 1) the current state s , 2) program inputs x_y , 3) parameters x_p , and 4) program variables x_v . Based on computations over these identifiers, the transition function returns the next state. Figure 4 has a list of operators that often appear in transition functions, such as arithmetic and trigonometric functions, but the list is *not* exhaustive. As a concrete example of a transition function written in repairable form, Figure 3a shows the transition function for the running example: it branches on the current state (s) and returns the next state. The crux of the transition function are the conditions that determine when the transition from GoTo to Kick occur.

A parameter map (P) specifies the parameters of a transition function. Figure 3b shows the parameter map of our example before running SRTR. The output of SRTR will be a set of adjustments to this parameter map.

An execution trace is a sequence of *trace elements* τ_t . A trace element records the values of sensor inputs, program variables, and the state at the start of time-step t . Finally, a user-provided *correction* (c_t) specifies the expected state at the end of time-step t . In our example, the attacker should have transitioned to the Kick state after $t = 5$.

Figure 3b shows the trace element and a user correction for the running example: since the correction c_5 refers only to the time-step $t = 5$, only the relevant trace element τ_5 is shown.

4.2 Residual Transition Functions

The goal of SRTR is to adjust the parameters such that for each correction (c_t) the transition function produces the corrected next-state instead of the actual state recorded at time $t + 1$. SRTR first simplifies the problem by specializing the transition function using the state, variables, and inputs recorded at time t . We call this simplified function the *residual transition function*. Figure 3c shows the residual transition function for the correction at $t = 5$. Since the state at this time-step ($\tau_5.state$) is GoTo, the residual transition function only has the code from the branch that handles this case (i.e., the code from lines 3–13). Furthermore, we substitute the input and variable identifiers with concrete values from the trace element and simplify expressions as much as possible. Therefore, the only identifiers that remain are parameters. This approach is known as *partial evaluation* [Jones *et al.*, 1993]. At a later step, SRTR translates this residual transition function into a formula for an SMT solver.

A potential problem with this approach is that SMT solvers do not have decision procedures that support trigonometric functions, which occur frequently in RSMs. Our example also uses trigonometric functions in several expressions. Fortunately, most of these trigonometric functions are applied to inputs and variables, thus they are substituted with concrete values in the residual. For example, line 6 calculates $\sin(robotAng_y)$ and $\cos(robotAng_y)$, but $robotAng_y$ is an input. Thus, the residual substitutes the identifier with its value

<p>Transition function (T)</p> <pre> 1 if (s == "START") { 2 return "GOTO"; 3 } else if (s == "GOTO") { 4 relLoc := ballLoc_y - robotLoc_y; 5 aimErr := AngleMod(targetAng_y - robotAng_y); 6 robotYAxis := ⟨sin(robotAng_y), -cos(robotAng_y)⟩; 7 relLocY := robotYAxis · relLoc; 8 maxYLoc := maxDist_p · sin(viewAng_p); 9 if (aimErr < aimMargin ∧ relLoc < maxDist_p ∧ 10 relLocY < maxYLoc) { 11 time_y > lastKick_v + kickTimeout_p } { 12 return "KICK"; 13 } else return "GOTO"; 14 } else if (s == "KICK" ∧ 15 timeInKick_v > kickTimeout_p) { 16 return "END"; 17 } else return "KICK"; </pre> <p>(a) A simple RSM and its transition function.</p>	<p>Parameter map (P)</p> $P = \langle \text{aimMargin}_p \mapsto \pi/50, \text{maxDist}_p \mapsto 80, \text{viewAng}_p \mapsto \pi/6, \text{kickTimeout}_p \mapsto 2 \rangle$ <hr/> <p>Trace element (τ_5)</p> $\tau_5.\text{in} = \langle \text{ballLoc}_y \mapsto \langle 30, 40 \rangle, \text{robotLoc}_y \mapsto \langle 0, 0 \rangle, \text{robotAng}_y \mapsto 0, \text{targetAng}_y \mapsto \pi/60, \text{time}_y \mapsto 5 \rangle$ $\tau_5.\text{vars} = \langle \text{lastKick}_v \mapsto 2, \text{timeInKick}_v \mapsto 0 \rangle$ $\tau_5.\text{state} = \text{"GOTO"}$ <hr/> <p>Trace (\mathcal{R})</p> $\mathcal{R} = \langle \dots \tau_5 \dots \rangle$ <hr/> <p>Correction (c_5)</p> $c_5 ::= s_6 \mapsto \text{"KICK"}$ <p>(b) Inputs to SRTR.</p>	<p>Repairable and unrepairable parameters</p> $\text{Rep}(T) = \{ \text{aimMargin}_p, \text{maxDist}_p, \text{kickTimeout}_p \}$ $\text{Unrep}(T) = \{ \text{viewAng}_p \}$ <hr/> <p>Result of MakeResidual(T, τ_5, P)</p> <pre> 1 if (π/60 < aimMargin_p ∧ 50 < maxDist_p ∧ 2 40 < maxDist_p · 0.5 ∧ 5 > 2 + kickTimeout_p) { 3 return "KICK"; 4 } else return "GOTO"; </pre> <hr/> <p>Result of CorrectOne(T, τ_5, P, c_5):</p> $\phi = \exists \delta^1, \delta^2, \delta^3 : \pi/60 < \pi/50 + \delta^1 \wedge 50 < 80 + \delta^2 \wedge 40 < (80 + \delta^2) \cdot 0.5 \wedge 5 > 2 + (2 + \delta^3)$ <hr/> <p>Result of CorrectAll($T, P, \mathcal{R}, \{c_5\}$):</p> $\Phi = \exists \delta^1, \delta^2, \delta^3, w^1 : w^1 = H \vee (w^1 = 0 \wedge \phi)$ <hr/> <p>Result of SRTR($T, P, \mathcal{R}, \{c_5\}$) for $H = 1$:</p> $\arg \min_{w^1, \delta^1, \dots, \delta^3} w^1 + \ \delta^1\ + \ \delta^2\ + \ \delta^3\ \text{ constrained by } \Phi$ $= \langle w^1 = 0, \delta^1 \mapsto 0, \delta^2 \mapsto 0.5, \delta^3 \mapsto 0 \rangle$ <p>(c) Each step of the SRTR algorithm.</p>
---	--	---

Figure 3: SRTR applied to a simplified robot soccer attacker with a single correction.

<p>Unary Operators</p> $op_1 ::= - \mid \sin \mid \cos \mid \dots$ <p>Expressions</p> $e ::= k \quad \text{Constants}$ $s \quad \text{State}$ $x_v \quad \text{Variables}$ $x_y \quad \text{Inputs}$ $x_p \quad \text{Parameters}$ $op_1(e)$ $e_1 \ op_2 \ e_2$ <p>Transition Functions</p> $T ::= \{ m_1; \dots; m_n \}$ <hr/> <p>Trace Elements</p> $\tau_t ::= \{ \text{in} = \langle \bigvee_{i=1}^m x_y^i \mapsto k^i \rangle, \text{vars} = \langle \bigvee_{j=1}^n x_v^j \mapsto k'^j \rangle, \text{state} = s_t \mapsto k'' \}$	<p>Binary Operators</p> $op_2 ::= + \mid - \mid * \mid > \mid \dots$ <p>Statements</p> $m ::= \text{return } s;$ $x_v := e;$ $\text{if } (e) \ m_1 \ \text{else } m_2$ $\{ m_1 \dots m_n \}$ <p>Parameter Maps</p> $P ::= \langle x_p^1 \mapsto k^1 \dots x_p^n \mapsto k^n \rangle$ <p>Traces</p> $\mathcal{R} ::= [\tau_1 \dots \tau_n]$ <p>Corrections</p> $c_t ::= s \in S$
--	--

Figure 4: Syntax of transition functions, traces, and corrections.

from the trace element ($\tau_5.\text{robotAng}_y = 0$) and simplifies the trigonometric expressions. In contrast, line 8 applies a trigonometric function to a parameter ($\sin(\text{viewAng}_p)$). This makes viewAng_p an *unrepairable parameter* that cannot appear in the residual transition function. SRTR substitutes unrepairable parameters with their concrete values from the parameter map.

In general, the MakeResidual function of SRTR (lines 11–15 in Figure 5) takes a transition function (T), a trace element (τ_t), and a parameter map (P) and produces a residual transition function by partially evaluating the transition function with respect to the trace element and the unrepairable parameters. We use a simple dataflow analysis to calculate the unrepairable parameters ($\text{Unrep}(T)$) and a canonical partial evaluator (Peval) [Jones *et al.*, 1993].

4.3 Transition Repair as a MaxSMT Problem

Given the procedure for calculating residual transition functions, SRTR proceeds in three steps. 1) It translates each correction c_i into an independent formula ϕ_i . A solution to ϕ_i corresponds to parameter adjustments that satisfy the correction c_i . Note however that no solution exists if c_i cannot be satisfied. 2) It combines the formulas $\phi_{1..n}$ for all cor-

rections $c_{1..n}$ from the previous step into a single formula Φ with independent penalties w^i for each sub-formula ϕ_i . A solution to Φ corresponds to parameter adjustments that satisfy a subset of the corrections. Any unsatisfiable corrections incur a penalty. 3) Finally, it formulates a MaxSMT problem that minimizes the magnitude of adjustments ($\|\delta^j\|$) to the parameters, and the penalty (w^i) of violated sub-formulas.

The CorrectOne function transforms a single correction into a formula. This function 1) calculates the residual transition function (Figure 5, line 18), 2) gets the repairable parameters (line 19), and 3) produces a formula (line 20) with a variable for each repairable parameter. In our running example, the transition function has four parameters, but, as explained in the previous section, the residual has only three parameters since viewAng_p is unrepairable. Therefore, the formula that corresponds to this residual (Figure 3c) has three variables (δ^1 , δ^2 , and δ^3). Moreover, since the correction (c_5) requires the next-state to be Kick, which only occurs when the residual takes the true-branch (line 3 of the residual), the body of the formula is equivalent to the conditional expression (lines 1–2), but with each parameter replaced by the sum of its concrete value (from P) and its adjustment (a δ). For example, the formula replaces aimMargin_p by $\pi/50 + \delta^1$. Therefore, when $\delta^1 = 0$, the parameter is unchanged.

The CorrectAll function supports multiple corrections and uses CorrectOne as a subroutine. The function iteratively builds a conjunctive formula Φ , where each clause has a distinct penalty w^i and two mutually exclusive cases: either $w^i = 0$ thus the clause has no penalty and the adjustments to the parameters satisfy the i th correction (line 27); or a penalty is incurred ($w^i = H$) and the i th correction is violated. Thus $H \in \mathbb{R}^+$ is a hyperparameter that induces a tradeoff between satisfying more corrections *vs.* minimizing the magnitude of the adjustments: large values of H satisfy more corrections with larger adjustments, whereas small values of H satisfy fewer corrections with smaller adjustments. The final formula has m real-valued variables δ^i for adjustments to the corresponding m repairable parameters x_p^i , and n discrete variables w^j that represent the penalty of violating


```

1 // Takes a transition function  $T$ , and returns a partially evaluated residual transition
2 // function  $T'$  by eliminating identifiers  $x^i$  using their values  $k^i$ .
3 def Peval( $T, x^1 \mapsto k^1 \dots x^n \mapsto k^n$ );
4
5 // Returns the list of repairable parameters of the transition function  $T$ 
6 def Rep( $T$ );
7
8 // Returns the list of unrepairable parameters of the transition function  $T$ 
9 def Unrep( $T$ );
10
11 def Residual( $T, \tau_t, P$ ):
12   {in =  $\langle \bigvee_{i=1}^l x_p^i \mapsto k^i \rangle$ , vars =  $\langle \bigvee_{j=1}^m x_v^j \mapsto k'^j \rangle$ , state =  $s_t \mapsto k''_s$ } :=  $\tau_t$ 
13   { $x_p^1, \dots, x_p^n$ } = Unrep( $T$ )
14    $T' :=$  Peval( $T, \bigvee_{i=1}^l x_p^i \mapsto k^i, \bigvee_{j=1}^m x_v^j \mapsto k'^j, \bigvee_{k=1}^n x_p^k \mapsto P(x_p^k), s_t \mapsto k''_s$ )
15   return  $T'$ 
16
17 def CorrectOne( $T, \tau_t, P, c_t$ ):
18    $T' :=$  Residual( $T, \tau_t$ )
19   { $x_p^1, \dots, x_p^m$ } := Rep( $T$ )
20   return  $\exists \delta^1, \dots, \delta^m : c_t = T'(s_1, x_p^1 + \delta^1, \dots, x_p^m + \delta^m)$ 
21
22 def CorrectAll( $T, P, \mathcal{R}, \{c_t^1, \dots, c_t^n\}$ ):
23   { $x_p^1, \dots, x_p^m$ } = Rep( $T$ )
24    $\Phi =$  true
25   for  $i \in [1 \dots n]$ :
26      $\exists \delta^1, \dots, \delta^m : \phi_i =$  CorrectOne( $T, \mathcal{R}[i], P, c_t^i$ )
27      $\Phi = \Phi \wedge (w^i = H \vee (w^i = 0 \wedge \phi_i))$ 
28   return  $\exists \delta^1, \dots, \delta^m, w^1, \dots, w^n : \Phi$ 
29
30 def SRTR( $T, P, \mathcal{R}, \{c_t^1, \dots, c_t^n\}$ ):
31   assert(CorrectAll( $T, P, \mathcal{R}, \{c_t^1, \dots, c_t^n\}$ ))
32   minimize( $\sum_{i=1}^m w^i + \sum_{j=1}^n \|\delta^j\|$ )
33   { $x_p^1, \dots, x_p^m$ } = Rep( $T$ )
34   return [ $x_p^1 \mapsto P(x_p^1) + \delta^1, \dots, x_p^m \mapsto P(x_p^m) + \delta^m$ ]
    
```

Figure 5: The core SRTR algorithm.

formula ϕ_j corresponding to correction c_t^j .

Our example (Figure 3c) has one correction and three repairable parameters. Therefore, CorrectAll produces a formula with four variables: a single penalty (w^1) and the three adjustments discussed above (δ^1 , δ^2 , and δ^3). The formula is a single exclusive-or: either the penalty is zero and formula is equivalent to the result of CorrectOne or the penalty is one and the result of CorrectOne is ignored.

Finally, the SRTR function uses CorrectAll as a subroutine and invokes the MaxSMT solver. This function 1) adds the assertion returned by CorrectAll to the solver, 2) directs the solver to minimize the sum of penalties and the sum of the magnitude of parameter changes (line 32 in Figure 5), and 3) returns a map from repairable parameter names to their new values. In our example, the solver calculates that the minimum-cost solution has $\delta^2 = 0.5$ with other variables set to zero. *i.e.*, we can satisfy the correction by adjusting maxDist_p from 80 to 80.5.

In summary, SRTR adjusts parameters to satisfy user-provided corrections. It is not always possible to find an adjustment that satisfies all corrections. Moreover, there is a tradeoff between making larger adjustments and satisfying more corrections. Therefore, SRTR uses a MaxSMT solver to formulate the parameter adjustment problem. A limitation of this approach is that a parameter may appear in a context cannot be expressed as a formula for the MaxSMT solver. We

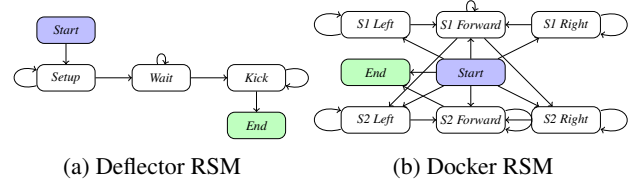


Figure 6: RSMs used for experiments

Method	Success Rate (%)	CPU Time
Initial Parameters	44	—
Exhaustive Search	89	1,300 hr
SRTR	89	10 ms

Table 1: Success rate and CPU time compared to exhaustive search.

use a simple analysis to determine which parameters cannot be repaired.

5 Evaluation

We evaluate SRTR in four ways. 1) We compare SRTR to an exhaustive search; 2) We show how the number of corrections affect RSM performance and that SRTR requires only a small number of corrections to perform well; 3) Using three RSMs, we show that SRTR does not over-fit and performs well in new scenarios; and 4) We use SRTR to improve the performance of a real-world robot.

We present three RSMs in this section and measure their success rates as follows. The *attacker* (Figure 1a) fills the main offensive role in robot soccer. Its success rate is the fraction of the test scenarios where it successfully kicks the ball into the goal. The *deflector* (Figure 6a) plays a supporting role in robot soccer, performing one-touch passing [Bruce *et al.*, 2008]. Its success rate is the fraction of the test scenarios where it successfully deflects the ball. The *docker* (Figure 6b) is a non-soccer behavior which drives a differential drive robot to line up and dock with a charging station. Its success rate is the fraction of the test scenarios where it successfully docks with the charging station.

5.1 Comparison To Exhaustive Search

Using the attacker, we compare SRTR to an exhaustive search to show that 1) SRTR is dramatically faster and 2) the adjustments found by SRTR are as good as those found by exhaustive search. To limit the cost of exhaustive search, the experiment only repairs the six parameters that affect transitions into the Kick state; we bound the search space by the physical limits of the parameters; and we discretize the resulting hypercube in parameter space. We evaluate each parameter set using 13 simulated positions and manually specify if the position should transition to the Kick state.

We evaluate the initial parameter values, the SRTR-adjusted parameters, and the parameters found by exhaustive search on 20,000 randomly generated scenarios. Table 1 reports the success rate and running time of each approach. SRTR and exhaustive search achieve a comparable success rate. However, SRTR completes in 10 ms whereas exhaustive search takes 1,300 CPU hours (using 100 cores).

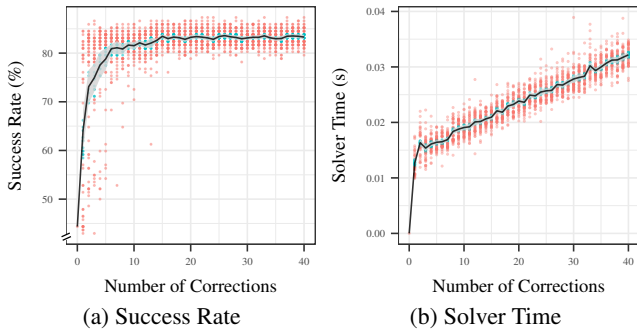


Figure 7: Scaling of SRTR with number of corrections. We report the mean as a line, the 99% confidence interval in grey, inliers in blue, and outliers in red. Darker points represent more occurrences.

5.2 Scalability

Using the attacker, we now show how the number of corrections affects SRTR solver time and success rate. Starting with the same set of initial parameters used in Section 5.1, we first create a training dataset by simulating the attacker in 40 randomly generated scenarios. For each scenario, we provide one correction, thus we have a training set of 40 corrections. Each trial applies SRTR to a subset of these corrections and evaluates the success rate using 150 random test scenarios. We repeat this procedure for each number of corrections $N \in [1, 40]$ with 50 randomly chosen subsets of the training set (*i.e.*, 300,000 total trials). Figure 7a shows how the number of corrections affects the success rate. It is possible for a single informative correction to dramatically increase the success rate, or for a particularly under-informative correction to have little effect. Therefore we also report the mean success rate and show the 99% confidence interval in gray. Success rate increases with the number of corrections, and with 23 corrections, the attacker reaches a peak success rate of 87%. Figure 7b shows that the solver time increases linearly with the number of corrections, and remains less than 0.04s with 40 corrections.

5.3 Generalizability

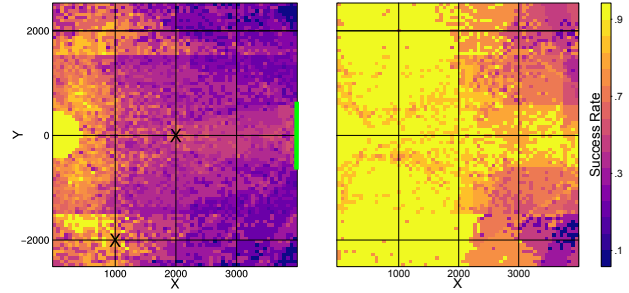
Since SRTR uses a handful of user-provided corrections to adjust parameters, there is a risk that it may over-fit and underperform in new scenarios. We use three RSMs to show that 1) SRTR-adjusted parameters generalize to new scenarios and that 2) SRTR outperforms a domain-expert who has 30 minutes to manually adjust parameters.

Table 2 summarizes the results of this experiment. We evaluate the success rate of the Attacker, Deflector, and Docker RSMs on a test dataset with several thousand test scenarios each. The baseline parameters that we use for these RSMs have a low success rate. We give a domain expert complete access to the RSM code (*i.e.*, the transition and emission functions), and subsequently our simulator for 30 mins. In that time, the expert is able to dramatically increase the success rate of the Deflector, but has minimal impact on the success rate of the Attacker and the Docker. Finally, we apply SRTR using a handful of corrections and the baseline parameters. The SRTR-adjusted parameters perform significantly better than the baseline and domain-expert parameters.

The heat maps in Figure 8 illustrates how parameters found

RSM	Params	SRTR Corrections	Tests	Success Rates (%)		
				Baseline	Expert	SRTR
Attacker	12	2	57,600	42	44	89
Deflector	5	3	16,776	1	65	80
Docker	9	3	5,000	0	0	100

Table 2: Success rates for baseline, expert, and SRTR parameters.



(a) Expert-adjusted parameters. (b) SRTR-adjusted parameters.

Figure 8: Attacker success rate with respect to different initial ball positions. The corrections are marked with a cross.

by the domain-expert, and by SRTR generalize to novel scenarios with the Attacker. In both heat maps, the goal is the green bar and the initial position of the Attacker is at the origin. Each coordinate corresponds to an initial position of the ball and for each position we set the ball’s initial velocity to 12 uniformly distributed angles. With the expert-adjusted parameters, the Attacker performs well when the ball starts in its immediate vicinity, but performs poorly otherwise. However, with SRTR-adjusted parameters, the Attacker is able to catch or intercept the ball from most positions on the field. For this result, we required only two corrections and the cross-marks in the figure show the initial position of the ball for both corrections. Therefore, although SRTR only adjusted parameters to account for these two corrections, the result generalized to many other positions on the field.

5.4 Case Study: SRTR In The Real World

To evaluate SRTR in the real world, we follow the same procedure that experts use (summarized in Table 3): we develop the Attacker in a simulator, we adjust parameters until it performs well in simulation, and then we find that it performs poorly in the real world. To evaluate the success rate of the attacker in the real world, we start the ball from 18 positions on the field and repeat each position five times with the same velocity (*i.e.*, 90 trials). The parameters from simulation have a 25% success rate. Using the execution logs of this experiment, we apply SRTR with three corrections. The adjusted parameters increase the success rate to 73%. In practice, an expert would iteratively adjust parameters, so we apply SRTR again with 2 more corrections, which increases the success rate to 86%. Finally, our group has an attacker that we tested and optimized extensively for RoboCup 2017, where it was part of a team that won the lower bracket. This *Competition Attacker* has additional states to handle special cases that do not occur in our tests. On our tests, the *Competition Attacker*’s success rate is 76%. Therefore, with two iterations of SRTR, the simpler Attacker actually outperforms the *Competition Attacker* in typical, real-world scenarios.

Trial	Success Rate (%)
Competition Attacker	75
Parameters from Simulation	24
Real World SRTR Tuning 1	73
Real World SRTR Tuning 2	85

Table 3: Attacker success rates on a real robot.

6 Conclusion

This paper presents SMT-based Robot Transition Repair (SRTR), which is a semi-automatic white-box approach for adjusting the transition parameters of Robot State Machines. We demonstrate that SRTR: 1) increases success rate for multiple behaviors; 2) finds new parameters quickly using a small number of annotations; 3) produces solutions which generalize well to novel situations; and 4) improves performance in a real world robot soccer application.

Acknowledgements

This work is supported in part by AFRL and DARPA agreement #FA8750-16-2-0042, and by NSF grants CCF-1717636 and CNS-1413985.

References

- [Abbeel and Ng, 2011] Pieter Abbeel and Andrew Y Ng. Inverse reinforcement learning. In *Encyclopedia of machine learning*, pages 554–558. Springer, 2011.
- [Argall *et al.*, 2009] Brenna D. Argall, Sonia Chernova, Manuela Veloso, and Brett Browning. A Survey of Robot Learning From Demonstration. *Robotics and Autonomous Systems*, pages 469 – 483, 2009.
- [Bai and Russell, 2017] Aijun Bai and Stuart Russell. Efficient Reinforcement Learning with Hierarchies of Machines by Leveraging Internal Transitions. In *ICRA*, 2017.
- [Biswas *et al.*, 2014] Biswas, Mendoza, Zhu, Choi, Klee, and Veloso. Opponent-driven planning and execution for pass, attack, and defense in a multi-robot soccer team. In *AAMAS*, pages 493–500, 2014.
- [Bjørner *et al.*, 2015] Nikolaj Bjørner, Anh-Dung Phan, and Lars Fleckenstein. *vZ*-an optimizing SMT solver. In *International Conference on Tools and Algorithms for the Construction and Analysis of Systems*, pages 194–199, 2015.
- [Bruce *et al.*, 2008] James Bruce, Stefan Zickler, Mike Licitra, and Manuela Veloso. CMDragons: Dynamic Passing and Strategy on a Champion Robot Soccer Team. In *ICRA*, pages 4074–4079, 2008.
- [Cano *et al.*, 2016] José Cano, Alejandro Bordallo, Vijay Nagarajan, Subramanian Ramamoorthy, and Sethu Vijayakumar. Automatic Configuration of ROS Applications for Near-Optimal Performance. In *IROS*, pages 2217–2223, 2016.
- [Dantam *et al.*, 2013] Neil Dantam, Ayonga Hereid, Aaron Ames, and Mike Stilman. Correct Software Synthesis for Stable Speed-Controlled Robotic Walking. In *RSS*, 2013.
- [Devlin *et al.*, 2017] Jacob Devlin, Jonathan Uesato, Surya Bhupatiraju, Rishabh Singh, Abdelrahman Mohammad, and Pushmeet Kohli. RobustFill: Neural program learning under noisy I/O. In *ICML*, 2017.
- [Gulwani, 2011] Sumit Gulwani. Automating string processing in spreadsheets using input-output examples. In *ACM SIGPLAN Notices*, pages 317–330. ACM, 2011.
- [Holtz and Biswas, 2017] Jarrett Holtz and Joydeep Biswas. Automatic Extrinsic Calibration of Depth Sensors with Ambiguous Environments and Restricted Motion. In *IROS*, pages 2235–2240, 2017.
- [Jones *et al.*, 1993] Neil D. Jones, Carsten K. Goman, and Peter Sestoft. *Partial Evaluation and Automatic Program Generation*. Prentice-Hall, Inc., 1993.
- [Kalmár-Nagy *et al.*, 2004] Tamás Kalmár-Nagy, Raffaello D’Andrea, and Pritam Ganguly. Near-Optimal Dynamic Trajectory Generation and Control of an Omnidirectional Vehicle. *Robotics and Autonomous Systems*, pages 47–64, 2004.
- [Kamar, 2016] Ece Kamar. Directions in Hybrid Intelligence: Complementing AI Systems with Human Intelligence. In *ICRA*, 2016.
- [Mechtaev *et al.*, 2015] Sergey Mechtaev, Jooyong Yi, and Abhik Roychoudhury. DirectFix: Looking for Simple Program Repairs. In *ICSE*, pages 448–458, 2015.
- [Meriçli *et al.*, 2012] Çetin Meriçli, Manuela Veloso, and H Levent Akin. Multi-resolution corrective demonstration for efficient task execution and refinement. *International Journal of Social Robotics*, pages 423–435, 2012.
- [Merici *et al.*, 2014] Cetin Mericli, Steven D. Klee, Jack Papparian, and Manuela Veloso. An Interactive Approach for Situated Task Specification Through Verbal Instructions. In *AAMAS*, pages 1069–1076, 2014.
- [Miyazawa *et al.*, 2017] Alvaro Miyazawa, Pedro Ribeiro, Wei Li, Ana Cavalcanti, and Jon Timmis. Automatic property checking of robotic applications. In *IROS*, pages 3869–3876, 2017.
- [Nashed and Biswas, 2018] Samer Nashed and Joydeep Biswas. Human-in-the-Loop SLAM. In *AAAI*, 2018.
- [Nedunuri *et al.*, 2014] Srinivas Nedunuri, Sailesh Prabhu, Mark Moll, Swarat Chaudhuri, and Lydia E. Kavraki. SMT-based synthesis of integrated task and motion plans from plan outlines. In *ICRA*, pages 655–662, 2014.
- [Reiter, 1987] Raymond Reiter. A Theory of Diagnosis from First Principles. *Artificial Intelligence*, pages 57–95, 1987.
- [Weiss *et al.*, 2017] Aaron Weiss, Arjun Guha, and Yuriy Brun. Tortoise: Interactive System Configuration Repair. In *ASE*, pages 625–636, 2017.
- [Wong *et al.*, 2014] Kai Weng Wong, Rüdiger Ehlers, and Hadas Kress-Gazit. Correct High-level Robot Behavior in Environments with Unexpected Events. In *RSS*, 2014.