Statistical Model Checking for Swarms – Input Attribution

Input Attribution – The "Why" of SMC

Statistical Model Checking (SMC) provides an estimate on the probability $P[\mathcal{M}\models\Phi]$ that a predicate Φ in a model $\mathcal{M}$ is satisfied, but does not address why a particular result was obtained. The goal of Input Attribution (IA) is to use machine learning techniques to synthesize an explanation for an SMC result in terms of the inputs. IA for SMC can be thought of as analogous to the counter-example in traditional model checking.

A good Input Attribution has the following properties:

1. Describes relationship that actually exists in data
2. Is presented in a way that is quantitative and understandable
3. Gives investigator new insights
4. Is resilient to randomness in the system

Example Scenario

Let (x_p,y_p) and (x_e,y_e) be random initial positions for a pursuer and an evader, respectively. The goal of the evader is to make it to one of several designated safe zones before it is caught by the pursuer. The SMC problem is to calculate the probability that the evader will escape. Intuitively, the probability of escape for the evader will depend on the initial distance between the pursuer and the evader, but can we synthesize this relationship purely from the SMC trials?

Approach – Logistic Regression

Logistic Regression (LR) is a regression model with a Boolean response variable based on the logistic function. A model $L{:}\vec{x}\rightarrow\mathcal{R}$ is generated from a set if input vectors $\vec{x}_i$ and corresponding Boolean responses ϕ_i. $L(\vec{x}_i)$ represents the log of the "odds" that the response ϕ_i is 1. The logistic function maps the log odds to a probability. The LR model is a linear function of the input variables with the form:

$$L{:}x\rightarrow\beta_0+\beta_1 x_1+\beta_2 x_2+\ldots+\beta_N x_N$$

Each coefficient β_j represents the factor by which the logit (log odds) of $\mathcal{M}\models\Phi$ increases for each unit increase of x_j. However, not all input variables may be statistically significant. When calculating each coefficient β_j, a standard error $se(\beta_j)$ that can be used to calculate a "p-value" indicating the significance of each coefficient is also produced. P-values greater than about 0.05 indicate that a particular input variable is not significant. The generated input attribution is formed from the β_j terms that are considered statistically significant.

Non-Linear Input Attribution

By expanding the Logistic model to include second order polynomial terms as:

$$L{:}\{\forall j{:}x_j,x_j\text{^}2\}\cup\{\forall j,k{:}x_j\,x_k\}\rightarrow\mathcal{R}$$

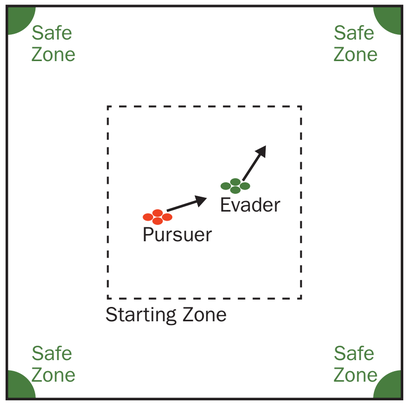
It is possible to discover more complex relationships among the input variables. After filtering terms that are not statistically significant, approximate factoring can be applied to pairs of terms to present the result in more human-readable form.

Validation

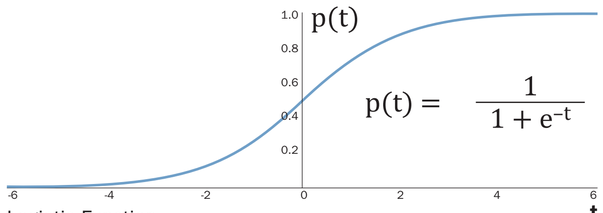
Even though LR analysis may indicate statistical significance on one or more variables, the overall model must have a good fit to the data before an input attribution can be accepted. We use the AUC (Area Under Curve) of an ROC (Radar Operating Characteristic) analysis as a metric. Five-fold cross validation is performed and the average AUC is used. AUC represents the probability $P[L(x_{SAT})>L(x_{UNSAT})]$ where x_{SAT} is an arbitrary satisfying input (ϕ=1) and x_{UNSAT} is an arbitrary unsatisfying input (ϕ=0). An AUC of 0.5 indicates the model is no better than guessing, while an AUC of 1.0 is a perfect model.

Experimental Results

We conducted SMC trials of the pursuer/evader scenario shown above using the V-REP simulation environment. Trials where conducted on a set if six 20-core blade servers. A target relative error of 0.01 was used which resulted in 39,960 trials. The resulting "mission success" probability for the evader was 0.214. The LR analysis and input attribution was conducted using the R statistical system and resulted in the expression shown to the right.



Example Scenario – Pursuer/Evader



Logistic Function

$$t = \cdots + 1.01\,x_p^2 - 2.03x_ex_p + 1.02x_e^2 + \cdots$$

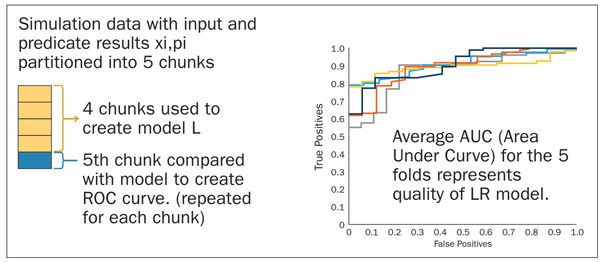
Look for approximate factorings

$$t = \cdots + 1.01\,(x_p - 1.01x_e)^2 + \cdots$$

Accepting approximation if error is small

$$t = \cdots + 1.01\,x_p^2 - 2.04x_ex_p + 1.03x_e^2 + \cdots$$

Approximate Factoring



5-Fold Cross Validation

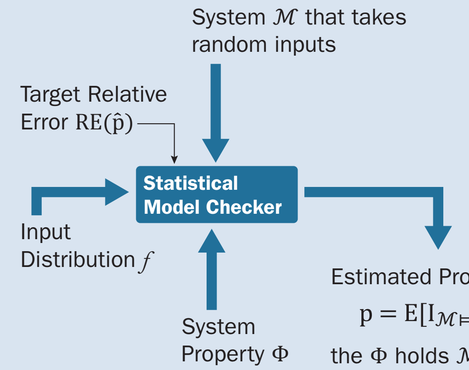
Name	ß	se(ß)	p-Value
x_ex_p	-0.124	0.0027	$< 10^{-4}$
y_ey_p	-0.122	0.0027	$< 10^{-4}$
x_e^2	0.060	0.0031	$< 10^{-4}$
y_e^2	0.056	0.0031	$< 10^{-4}$
x_p^2	0.056	0.0031	$< 10^{-4}$
y_p^2	0.056	0.0031	$< 10^{-4}$

Input Attribution Results

$$0.0602(x_e - 1.03x_p)^2 + 0.0561(y_e - 1.09y_p)^2$$

Factored Input Attribution

Statistical Model Checking (SMC) Basics



Relative Error

Measure of accuracy for a prediction

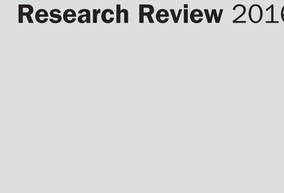
Defined as ratio of standard deviation to mean. For a probability estimate, the estimated relative error is:

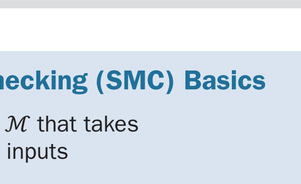
$$(RE) = \frac{\hat{\sigma}}{\hat{p}}$$

Number of samples to achieve a target relative error increases

- as target relative error decreases, or
- as estimated probability decreases

$$N \approx \frac{1}{p(RE)^2}$$





Conclusion

We applied SMC with Input Attribution to a pursuer/evader scenario. Intuitively we expected an Input Attribution indicating that increased initial distance between pursuer and evader should be correlated with improved chance of escape for the evader.

Distribution Statement A: Approved for Public Release; Distribution is Unlimited

Contact: Jeffery Hansen (jhansen@sei.cmu.edu)