Using Squeezeness to test from Finite State Machines

Manuel Núñez
(joint work with Alfredo Ibias and Rob Hierons)

Universidad Complutense de Madrid

CREST Information Theory and Software Testing

What's the meaning of Squeeziness?

What's the meaning of Squeeziness?

Honestly, until a couple of days ago, I had no clue!

What's the meaning of Squeeziness?

Honestly, until a couple of days ago, I had no clue!
I did some research for this talk....

Squeeziness is

- According to Wiktionary:

What's the meaning of Squeeziness?

Honestly, until a couple of days ago, I had no clue!
I did some research for this talk....

Squeeziness is

- According to Wiktionary:The quality of being squeezy....

What's the meaning of Squeeziness?

Honestly, until a couple of days ago, I had no clue!
I did some research for this talk....

Squeeziness is

- According to Wiktionary:The quality of being squeezy....
- Squeezy is: *flexible* or *causing compression*.

What's the meaning of Squeeziness?

Honestly, until a couple of days ago, I had no clue!
I did some research for this talk....

Squeeziness is

- According to Wiktionary:The quality of being squeezy....
- Squeezy is: *flexible* or *causing compression*.
- Related to an expression?

What's the meaning of Squeeziness?

Honestly, until a couple of days ago, I had no clue!
I did some research for this talk....

Squeeziness is

- According to Wiktionary:The quality of being squeezy....
- Squeezy is: *flexible* or *causing compression*.
- Related to an expression? easy peasy lemon squeezy.

What's the meaning of Squeeziness in Information Theory?

- D. Clark & R. Hierons. *Squeeziness: An information theoretic measure for avoiding fault masking*, IPL, 2012.
- It is a measure designed to quantify the likelihood of Failed Error Propagation.

What's the meaning of Squeeziness in Information Theory?

- D. Clark & R. Hierons. *Squeeziness: An information theoretic measure for avoiding fault masking*, IPL, 2012.
- It is a measure designed to quantify the likelihood of Failed Error Propagation.
- FEP happens when
  1. a faulty statement is executed during testing,

What's the meaning of Squeeziness in Information Theory?

- D. Clark & R. Hierons. *Squeeziness: An information theoretic measure for avoiding fault masking*, IPL, 2012.
- It is a measure designed to quantify the likelihood of Failed Error Propagation.
- FEP happens when
  1. a faulty statement is executed during testing,
  2. the fault *corrupts* the internal state of the SUT,

What's the meaning of Squeeziness in Information Theory?

- D. Clark & R. Hierons. *Squeeziness: An information theoretic measure for avoiding fault masking*, IPL, 2012.
- It is a measure designed to quantify the likelihood of Failed Error Propagation.
- FEP happens when
  1. a faulty statement is executed during testing,
  2. the fault *corrupts* the internal state of the SUT,
  3. but the expected output is observed.

What's the meaning of Squeeziness in Information Theory?

- D. Clark & R. Hierons. *Squeeziness: An information theoretic measure for avoiding fault masking*, IPL, 2012.
- It is a measure designed to quantify the likelihood of Failed Error Propagation.
- FEP happens when
  1. a faulty statement is executed during testing,
  2. the fault *corrupts* the internal state of the SUT,
  3. but the expected output is observed.

How bad is FEP?

- FEP can reduce testing effectiveness: we might fail to find a fault despite executing the faulty statement.
- Empirical studies show that many systems suffer from FEP.

What's this talk about?

The adaption of Squeeziness to a black box scenario.

This talk in a nutshell

Using Squeeziness to test from Finite State Machines

This talk in a nutshell

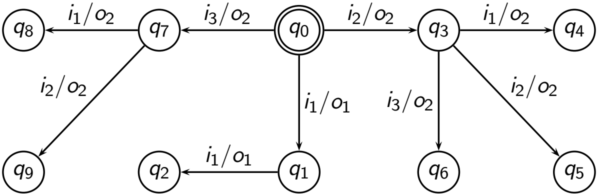
Using Squeeziness to test from Finite State Machines

Finite State Machines

Graphs with an initial state where transitions are labelled by a pair (input, output).

Finite State Machines

Graphs with an initial state where transitions are labelled by a pair (input, output).

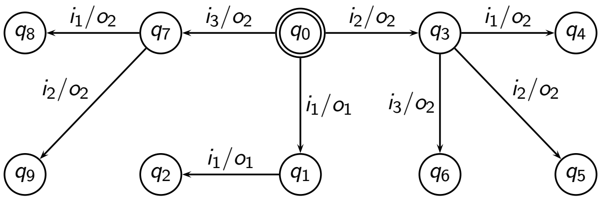


FSMs: assumptions

- FSMs are deterministic.
- FSMs representing SUTs are input-enabled.

FSMs as functions

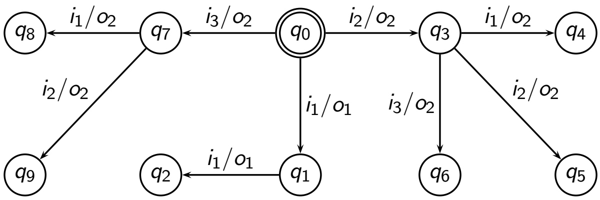
An FSM M can be seen as a function $f_M : \texttt{dom}_M \longrightarrow \texttt{image}_M$ such that for all $\alpha \in \texttt{dom}_M$ (sequence of inputs performed by M) $f_M(\alpha) = \beta$ (sequence of outputs observed after applying α).



$$
\begin{aligned}
f_M(i_1) &= o_1 \\
f_M(i_3) &= o_2 \\
f_M(i_3 i_1) &= o_2 o_2 \\
f_M(i_2 i_3) &= o_2 o_2
\end{aligned}
$$

FSMs as functions

An FSM M can be seen as a function $f_M : \texttt{dom}_M \longrightarrow \texttt{image}_M$ such that for all $\alpha \in \texttt{dom}_M$ (sequence of inputs performed by M) $f_M(\alpha) = \beta$ (sequence of outputs observed after applying α).



$$
\begin{aligned}
f_M(i_1) &= o_1 \\
f_M(i_3) &= o_2 \\
f_M(i_3 i_1) &= o_2 o_2 \\
f_M(i_2 i_3) &= o_2 o_2
\end{aligned}
$$

Collisions

α_1 and α_2 collide for M if $\alpha_1 \neq \alpha_2$ and $f_M(\alpha_1) = f_M(\alpha_2)$.

This talk in a nutshell

Using Squeeziness to test from Finite State Machines

Squeeziness as difference of entropies

Entropy of the random variable ξ_A

$$\mathcal{H}(\xi_A) = -\sum_{a \in A} \sigma_{\xi_A}(a) \cdot \log_2(\sigma_{\xi_A}(a))$$

If $f : A \longrightarrow B$ then Squeeziness of f is the loss of information after applying f to A: $\mathcal{H}(A) - \mathcal{H}(B)$.

Squeeziness for FSMs

We need to define how to group inputs & outputs. Two alternatives:

- A unique random variable for the whole set of inputs/outputs.
- A random variable for each length of sequences of inputs/outputs.

Squeeziness for FSMs

We need to define how to group inputs & outputs. Two alternatives:

- A unique random variable for the whole set of inputs/outputs.
- A random variable for each length of sequences of inputs/outputs.

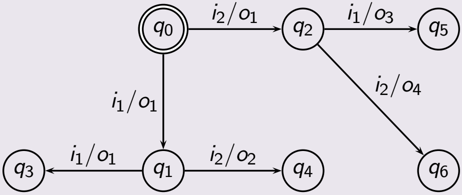
We choose the second one because it gives an incremental procedure to compute a sequence of *consecutive values* of Squeeziness.

Squeeziness as difference of entropies

Let FSM M, $k > 0$ and random variables $\xi_{\texttt{dom}_{M,k}}$ and $\xi_{\texttt{image}_{M,k}}$.

$$\texttt{Sq}_k(M) = \mathcal{H}(\xi_{\texttt{dom}_{M,k}}) - \mathcal{H}(\xi_{\texttt{image}_{M,k}})$$

Squeeziness is not monotonic



Squeeziness for $k = 1$ is $\log_2(2) = 1$ while for $k = 2$ is 0.
This is bad because we do not have an obvious stopping rule.

Squeeziness is null for bijective functions

If $f_{M,k}$ is bijective then $\mathtt{Sq}_k(M) = 0$.

Squeezines is null for bijective functions

If $f_{M,k}$ is bijective then $\mathtt{Sq}_k(M) = 0$.

Random variables for outputs are determined

Given FSM M, $k > 0$ and $\xi_{\mathtt{dom}_{M,k}}$, the probability distribution of $\xi_{\mathtt{image}_{M,k}}$ is completely determined.

$$\sigma_{\xi_{\mathtt{image}_{M,k}}}(\beta) = \sum_{\alpha \in f_M^{-1}(\beta)} \sigma_{\xi_{\mathtt{dom}_{M,k}}}(\alpha)$$

Maximum entropy principle

Maximum entropy is obtained with a uniform distribution $\xi_{\mathtt{dom}_{M,k}}$.

$$\mathtt{Sq}_k(M) = \frac{1}{|\mathtt{dom}_{M,k}|} \cdot \sum_{\beta \in \mathtt{image}_{M,k}} |f_M^{-1}(\beta)| \cdot \log_2(|f_M^{-1}(\beta)|)$$

Maximum entropy principle

Maximum entropy is obtained with a uniform distribution $\xi_{\mathtt{dom}_{M,k}}$.

$$\mathtt{Sq}_k(M) = \frac{1}{|\mathtt{dom}_{M,k}|} \cdot \sum_{\beta \in \mathtt{image}_{M,k}} |f_M^{-1}(\beta)| \cdot \log_2(|f_M^{-1}(\beta)|)$$

Maximum loss of information

Probability distribution maximising Squeeziness: uniformly distributed in the bigger inverse image of an element of the outputs β' and zero otherwise.

$$\sigma_{\xi_{\mathtt{dom}_{M,k}}}(\alpha) = \begin{cases} \frac{1}{|f_M^{-1}(\beta')|} & \text{if } \alpha \in f_M^{-1}(\beta') \\ 0 & \text{otherwise} \end{cases}$$

$$\mathtt{Sq}_k(M) = \log_2(|f_M^{-1}(\beta')|)$$

Probabilistic Squeeziness

We divide Squeeziness by its maximum value.

$$\mathtt{PSq}_k(M) = \frac{\mathcal{H}(\xi_{\mathtt{dom}_{M,k}}) - \mathcal{H}(\xi_{\mathtt{image}_{M,k}})}{\log_2(|f_M^{-1}(\beta')|)}$$

This talk in a nutshell

Using Squeeziness to test from Finite State Machines

Collisions and FEP

Let $m_i = |f_{M,k}^{-1}(\beta_i)|$ and $d = \sum_{i=1}^{n} m_i$. Assuming a uniform distribution, the probability of having a collision is:

$$\texttt{PColl}_k(M) = \sum_{i=1}^{n} \frac{m_i \cdot (m_i - 1)}{d \cdot (d-1)}$$

Relation between $\texttt{PColl}_k(M)$ and $\texttt{PSq}_k(M)$ is not monotonic

There exist M_1 and M_2 and $k > 0$ such that
$\texttt{PSq}_k(M_1) < \texttt{PSq}_k(M_2)$ but $\texttt{PColl}_k(M_1) > \texttt{PColl}_k(M_2)$.

Empirical Evaluation via simulations

Simulations to compute `PColl`, `PSq`, `Sq` assuming uniform distributions over the inputs (methodology similar to [CH12]).

- $d =$ size of the input space (ranging between 10^4 and $2 \cdot 10^9$).
- $m =$ maximum subdomain size (ranging between 10^2 and 10^4).

Empirical Evaluation via simulations

Simulations to compute `PColl`, `PSq`, `Sq` assuming uniform distributions over the inputs (methodology similar to [CH12]).

- d = size of the input space (ranging between 10^4 and $2 \cdot 10^9$).
- m = maximum subdomain size (ranging between 10^2 and 10^4).
- Pearson & Spearman Rank correlation coefficient between `PColl` and `PSq`/ `Sq`. Similar results.
- Strong correlation between `PColl` and `PSq`. Values greater than 0.96 for input sets with $5 \cdot 10^6$ or more elements.

Empirical Evaluation via simulations

Simulations to compute `PColl`, `PSq`, `Sq` assuming uniform distributions over the inputs (methodology similar to [CH12]).

- $d =$ size of the input space (ranging between 10^4 and $2 \cdot 10^9$).
- $m =$ maximum subdomain size (ranging between 10^2 and 10^4).
- Pearson & Spearman Rank correlation coefficient between `PColl` and `PSq`/ `Sq`. Similar results.
- Strong correlation between `PColl` and `PSq`. Values greater than 0.96 for input sets with $5 \cdot 10^6$ or more elements.
- Standard Squeeziness has a better correlation. Still, `PSq` can be more useful because it is easier to compare results from different machines and lengths of inputs.

Empirical Evaluation via FSMs: Squeeziness and fault location

- 50 randomly generated FSMs (between 25 and 50 states).
- For each FSM we computed $\mathtt{Sq}$ and $\mathtt{PSq}$ for all $1 \leq k \leq 25$.
- We generated 100 valid mutants of M presenting FEP.

Empirical Evaluation via FSMs: Squeeziness and fault location

- 50 randomly generated FSMs (between 25 and 50 states).
- For each FSM we computed $\mathtt{Sq}$ and $\mathtt{PSq}$ for all $1 \leq k \leq 25$.
- We generated 100 valid mutants of M presenting FEP.
- No correlation between where the fault is produced and the Squeeziness and Probabilistic Squeeziness obtained for the length of the input sequence reaching the mutated transition.
- Negative result. We tried something less ambitious.

Empirical Evaluation via FSMs: Squeeziness and probability of FEP

Instead of *predicting* where the fault was, we consider the probability of FEP.

- Same 50 randomly generated FSMs.
- We generated 50 valid mutants of M (with and without FEP).
- We computed the probability of FEP, `Sq` and `PSq` for length 25.
- $p(FEP) = \frac{\#\text{ tests reaching wrong state but generating correct output}}{\#\text{ tests reaching wrong state}}$.

Empirical Evaluation via FSMs: Squeeziness and probability of FEP

Instead of *predicting* where the fault was, we consider the probability of FEP.

- Same 50 randomly generated FSMs.
- We generated 50 valid mutants of M (with and without FEP).
- We computed the probability of FEP, $\mathtt{Sq}$ and $\mathtt{PSq}$ for length 25.
- $p(FEP) = \frac{\#\text{ tests reaching wrong state but generating correct output}}{\#\text{ tests reaching wrong state}}$.
- High correlations between probability of having FEP with sequences up to 25 and $\mathtt{Sq}$ and $\mathtt{PSq}$ for $k = 25$. All the values were greater than 0.75 and some close to 1.

This talk in a nutshell

Using Squeeziness to test from Finite State Machines

Application to testing

We may consider that $1-\mathtt{PSq}$ gives the *reliability* of tests: it represents the probability that a correct output indicates that no fault was executed.

Application to testing

We may consider that $1-\mathtt{PSq}$ gives the *reliability* of tests: it represents the probability that a correct output indicates that no fault was executed.

Before running tests, we may compute $\mathtt{PSq}$ for different values of k. We can choose a value of k such that $\mathtt{PSq}$ is low: this makes is less likely to have FEP.

Application to testing

We may consider that $1-\mathtt{PSq}$ gives the *reliability* of tests: it represents the probability that a correct output indicates that no fault was executed.

Before running tests, we may compute $\mathtt{PSq}$ for different values of k.

We can choose a value of k such that $\mathtt{PSq}$ is low: this makes is less likely to have FEP.

Finally, if we have $\mathtt{PSq}= 0$ for a certain k, we can use this length of tests as a checkpoint (but remember that we do not have monotonicity).

Conclusions and future work

- Squeeziness in a black-box framework.

Conclusions and future work

- Squeeziness in a black-box framework.
- No correlation between Squeeziness for k (length of tests) and faults at length $k-1$.
- Correlation between Squeeziness and probability of FEP.

Conclusions and future work

- Squeeziness in a black-box framework.
- No correlation between Squeeziness for k (length of tests) and faults at length $k-1$.
- Correlation between Squeeziness and probability of FEP.
- Future work: Consider observable FSMs and experiments on *real* FSMs.

THANKS FOR YOUR ATTENTION!!
Questions? Comments?