Discussion Part 1: Identifying internal stage compositions

Instrumentation

Participant
If I start with the instrumentation and one thing which comes immediately is in my mind is that you typically use or often use instrumentation, for instance to collect the coverage. So there is at least from my perspective, I'll just go ahead and do this.

Certain link between the instrumentation and the execution because once you do the instrumentation.
Well, you will be put into a situation where you are able to then get certain insights about the execution that you wouldn't get otherwise.

I also think there are at least one, but possibly also two other connections. And The thing is, if I draw a connection between the power schedule first, but also to the search strategy, I guess. And I'm thinking at the moment about AFLGo because in AFLGo, what you do is that you prioritize certain seeds and also the adaptive number of mutations that you perform in it. I think depending on how close the target is to the current input to the goal that we have to find.

Interviewer
I'm just wondering if it is it the property of the instrumentation that does this or is it the property of some initial seeds or some search strategy or some mutations that does this?

Participant
The thing is, what you do in aflgo is that you instrument the distance. So you have two compilation steps. In the first step you calculate you know graph. And then during the runtime to make it faster the distance calculation is actually done within the instrumentation.

So this then gives insights for the search strategy and the power schedule because they derive their values and I think it's possibly more of the power schedule. I'm not quite sure because so we might possibly delete the search strategy because I think

it's still doing the typical AFL thing. But in terms of how often you mutate a seed that is based on the distance calculations and these are instrumented.
So if you visit a basic block, then the value that you get, let's say in terms of fitness or whatever: this is based on the compile time instrumentation and this is also instrumented into the system under test.

I don't have anything really in terms of the initial seats or the mutations at the moment.

Initial Seeds

Participant
For me this it might be a bit more difficult.

Because if I'm being honest, I also remember having read the about the importance of the initial seeds.

But have I ever seen or thought about connections between the initial seeds and the other components terms of? Search strategy and power schedule.
To be honest, at the moment I don't have any.

Interviewer
Alright, so maybe I can ask you a question here so you can tell me if that is valid, right. So for initial seats, from my opinion, it's sometimes related to the mutations, right? So the mutants that you get from a good initial seed is what makes the further process of fuzzing perform better. Do you agree on this or do you think that's not at all related?

Participant
Yeah, this is correct. So depending on the quality of the mutants, derided from the initial seeds, you have to do fewer, let's say mutations, for example, to still get about input which is more or less semi-valid. So I would agree that there is a relation between those two.

And I mean, if we're thinking about this, it also definitely makes sense to think about in terms of the execution, I think because depending on what kind of seeds you provide also effects what kind of code you visit.

There could be also.
I was thinking about, of course, if you have depending on the set of seeds that you have, it could also influence what you later analyse further, but this wouldn't be the initial seeds as they would be produced. So to really pick out something you would have to need some sort of information and if you to get this information you first have to run them.

Search Strategy

Interviewer
So I think I would just go ahead and draw an arrow from search strategy to power schedule because you mentioned before.

Participant
The thing is typically, what you do try to do in the search strategy is to, well, select to see which you find promising, and this already gives you some sort of indication. You know how often you want to produce it. So let's say intuitively, if you pick a seed pretty early, you would also assume that you would produce many mutants of that. So there is, I think, some sort of correlation, let's say one can a positive correlation between how likely it is to pick a seed early and also how many mutants you produce of that. So they're interconnected. But in theory they are. If there's some reasoning, they could be also separated.

And I think similar thing could also possibly be in terms of the mutations that apply. So you might have a search strategy and the mutations: You can either do many mutations or a few mutations. And if the search strategy for instance selects a certain seed at a certain time point, It might also do this because of things, for example, that this seed has produced a certain behavior: for example, it has come quite close to target line of code. But it also knows that due to the small distance that we have, the kinds of mutations that we want to perform shouldn't be too drastic.

So it will then possibly influence the types of mutations that are performed but also the extent.

Let's say in terms of execution.
The search strategy directly influences what is executed, so only what is picked up by the search strategy as a seed is done, also executed by. This might be too obvious of in connection.

Then think at the moment in terms of instrumentation and initial seeds. I don't think there is one. So I would be done for the moment.

Power Schedule

Interviewer
Right, let's go to the power schedule.

Participant
To be honest, any link to the first three because it is unidirectional, so we will definitely have the flow from the search strategy to the power schedule, but typically not the other way around.

In terms of mutations?
Once again, there could be a certain link if we decide to do many.
So let's say we have a power schedule. Power schedule says well, we want to do, we wanted to produce many mutants.

Then this might also impact the extent of the mutations, for example,
If we want to do many or rather small mutations, however, there's still something I think we just possibly more really directed by the search strategy, I think the rather obvious link is that the power schedule, of course affects the executions, so the real the power schedule in effect tells you, well, how often do you then produce?

So this definitely effects how what kind of inputs we create, let's say what is effectively executed.

However, in terms of the power schedule to mutations.
I think this would be really something which is more possibly directed by the search strategy because the power schedule gets as input. I think something which and this type of input would also rather be guiding for the mutation. So I think I would leave it at that.

Mutations

Participant
I mean, of course the mutations impact what we execute as they directly influence that, do they impact the instrumentation? No, this is of course aftewards.
So we would skip that.

So the direct link between the mutations and the initial seeds, let's think of that. I mean, if you have a certain set of mutations that you want to execute.
It might impact the initial seeds that you are dealing with.
Let's say you have some sort of fuzzer which has domain knowledge. Then you need very specific seeds which are for instance well formed. Let's say you have, fuzzer which can deal with inputs which are based on a grammar. And you want to perform grammar based mutations that you then you would also need seeds which are well formed so that you can parse them and then thus later also mutate them in a wise way that you can work with. Otherwise this whole process does not work. So I mean, if I think that's correct, I could draw this connection.

In terms of the search strategy.
Does it do the mutations? Impact? What kinds of inputs you select?

We have a directional link from the mutations to the search strategy. So the types of mutations, can they affect the search strategy?

Interviewer
Maybe is it interesting to you to know how many times the mutants of from this particular mutation phase are selected?

Participant
This is a good point. Yeah, this makes absolute sense sometimes I think.
I've seen this I guess also sometimes in the same my work. You might try to do a certain type of mutation and you evaluate then the impact of the mutation.

And this then can indeed have an impact on what kind of inputs you would later select possibly. So let's say you have an input and you mutated it as couple of times. However, it did not produce the results that you wanted to have. This then would impact both what you would select and also how often. So it would definitely say we have links to both in this case.

Execution

Participant
OK, execution.
Yeah, I think so, in theory we have more or less to almost all the boxes any sort of. The thing is, why is it related to initial seeds? Well, you might have observed in previous campaigns that your initial seed set is insufficient to produce the behavior or analyze the behavior that you want, so definitely you will then adapt the initial seats or this might be one option.

Also the execution might tell you that your current search strategy might not be effective, so let's say you have a seat set which is there to cover all the general behaviors, however you notice that due to the reserve strategy, you'd tend to end up only fuzzing very specific kinds of functionalities that you do not want.

So in this case feedback that you would get from the execution would both adapt the search strategy, but once again, also the number of mutants that you would generate as well as possibly the types of mutations. So as again you might see that if you execute mutant or a particular type of mutation that you will not get the results that you would like to have and this could then impact what kinds of mutation types you do and how they you know what extent they end.

Discussion Part 2: Identifying visualization analysis tasks

Participant
So it's saying.
Uh.
Man translates how does that? OK, so we want to understand the instrumentation.
Umm.
Well, the first thing is of course to ask question, what effect does the So what is the goal of the instrumentation we would typically there are.

Are you writing it down, or should I just go ahead?

Interviewer
You can right now.

Instrumentation

P5.1 AT. What is the goal of the instrumentation?

Participant
I guess so. What is the goal of the instrumentation? This is the first thing.

P5.2 AT. What parts of the code are instrumented?

Participant
It's quite obvious. Typically we want to achieve something with that. The second thing is what does the instrumentation actually do.
What part of the code are instrumented? Because it could be.

P5.3 AT. How much additional code is added by the instrumentation?
P5.4 AT. How much overhead produced by the instrumentation (in terms of execution)? Possibly also in terms of compilation.

Participant
Code instrumented so it could be it's only part of the code.
But another thing which might be also interesting, not only what parts of the code

are instrumented, but also how much, let's say additional code as this would be the performance. And this possibly leads us also to another thing, how much overhead?

But it might be also interesting, at least in some scenarios.
in terms of compilation time?

If I remember the compilation in AFLGo also takes some considerable time.
So if we could also visualise it in some form and also consider it doing the comparison, this would be in helpful I think.

P5.5 AT. How does the instrumentation impact the metric that is used for the assignment of energy in the power schedule?

Participant
Speaking of grey-box, we have this relationship between the instrumentation and the power schedule. Is there any way for me to?

Interviewer
In you said when you were right. I'm putting this. I think you said how does instrumentation impact energy assignment.

Participant
In this particular case, you instrument within the binary.
A certain value which is then used for the fitness.

One thing one could possibly look at in certain sense is to see, how much the parts which are added by the instrumentation, really affect the number of let's say for example the fitness value or whatever other metric is.
How does the instrumentation impact the metric that is used for.
So yeah, assignment of energy.

Different things depending on if it's all in, for example, provides a certain.
Component only or how much of the fitness it really contributes. There was only different ways to look at this.

Initial Seeds

P5.6 AT. How does the "human-readable" version of the seed look like?

Participant
Well, one thing one could typically do is there are certain seats which can be really visualized so. For example, if we think of an image or something like that.
What is the visual representation of the seed? So how does the human readable.
I Will add quotation marks, but the seed looked like because I think of it in two terms. First of all, you have certain binary formats, like an image, which you can directly visualize or you have text format like an XML. So if you have an XML which is also clearly malformed. This would also provide you insights into the seeds. Which you wouldn't get if you, for example, would look only at the raw bites which make up the input in this case.

P5.7 AT. How do differences between seeds look like? E.g., with "diff" on text files.

Participant
One thing which might be interesting if in terms of diffs, if you could perform a binary diff, assuming that you have so how do you how would do differences between?

P5.8 AT. What parts of the initial seeds are mutated?

Participant
Instance or with diff on text files.
Alright then in terms of the Mutations and the execution. Well, I would find interesting in terms of this mutations and initials seeds: what parts of the initial seeds are mutated to get an idea, also especially then for the execution.

P5.9 AT. What parts of the code are covered by the initial seeds? Or possibly other behaviors (e.g., memory or CPU consumption).

So this is a bit related, but also then for the 4th question would be what let's say parts of the code are covered by the initial seeds. And not only covered, but also for

instance, let's say whatever possibly other behaviors like memory or CPU consumption?

Search Strategy

P5.10 AT. What kind of seeds are selected how often?

Participant
Through such a power schedule, what I would possibly find interesting in terms of visualization is if I have a queue to possibly get highlighted, let's say in a color coded manner or what kind of seeds are particularly selected often.

P5.11 AT. What kinds of seeds are selected when? E.g., in terms stages (and possibly time).

Search strategy also could impact the mutations on seed set because we might have a mutate. A search strategy would want wants to pick out specific ones.
So one thing which might be interesting, for instance, let's say we have a search strategy which changes depending on different let's say stages.
First, it does some sort of random fuzzing and then it more guided fuzzing. One thing which might be interesting to know would also be, I mean it's again more in terms of the power schedule, but, what kind of seeds are selected when. For instance, in terms of time, if it makes sense for us, or rather in terms of stages.
I guess, maybe even the more important, possibly time. It's interesting to anyone.

P5.12 AT. What are the correlations between the search strategy (or its current state) and the mutations (e.g. types and extent) which are performed?

Participant
Let's see search strategy and mutations. There is this connection between search strategy might in fact effect the extent of mutations that we have because it might do a certain stage, can we get some insights?

Interviewer
I think in the previous step you mentioned some correlation.

Participant
You would definitely have a correlation between what you pick

What is the correlation between the search strategy or its current status and the mutations which are performed.
And I think I would be then done for that.

Power Schedule

P5.13 AT. How many children are produced off a seed?

Participant
So let's have a look at the power schedule. So in terms of power schedule.
This now means again.
Yeah, we have a similar thing. Some sort of visualization on how often particular seed is selected and how many children are produced? in some sort of heat map possibly.

P5.14 AT. How does energy assignment vary over time per seed?

Participant
Could be also interesting to see. How it does change over time?
How? I mean, there could be certain feathers where or just generally where the. Whether inputs which are select first or put first into the queue are at some point they are not often selected, so it might if you also see it in some sort of visual sense that it really degrades, then you could then decide OK, maybe it might be time for some sort of strategy which just removes them from the queue because they do not add any benefits. So possibly also how many children are produced.
See it over time. How this evolves?

P5.15 AT. How many children overall?

Yeah, maybe one thing which could be also interesting is how many children or mutants are produced overall.

Interviewer
OK so I it looks like these are the questions for the mutations to me.

Participant
Well, the past schedule the sites are many children we produce, right?

P5.16 AT. How many children are produced off a seed over time?

Interviewer
Ah, right, right. OK. Sorry. My bad. OK, how many children are produced of a seat over time? OK.

Participant
Because it might sense not only to look at. How many are produced for a particular seed, but also just in terms of overall inputs that the fuzzer tends to generate? Maybe in the within a particular time frame could be that for some reason just wants to do a few executions. So this might be an interesting thing.

Mutations

P5.17 AT. What parts of the input are mutated?

Participant
OK, now all mutations definitely.
What parts of the input are mutated?

P5.18 AT. How does a mutation impact the following decisions of a fuzzer? : mutation – search strategy, mutation – power schedule, mutation – initial seeds

Participant
So the mutations could impact the search strategy as there might be definitely a link between the result of the mutation and how it impacts the other processing steps. So not only the search strategy, so maybe keep it more general. So how does the mutation impact the following decisions, let's say.

So let's say we have some sort of mutation and that either produces a result we like and then we see, OK, this is a shift for instance, the search strategy that certain inputs are prioritized more or else in terms of the power schedule that we generate more

inputs out of that and these then could be I think answered or better analyzed with the visualization guides or aids that we have already defined.

P5.19 AT. What impact does a mutation have on the coverage?

Participant
In terms of mutation and execution, I think we have once again a similar thing to the initial seeds. So assume we do some sort of mutation: does it cure some sort of new coverage and can we also visualize it to see this type of mutation that is sort of input.

So how? What's?
Want to impact does a mutation.

Interviewer
Something like mutation. How it is, how it performs in terms of coverage?

Participant
Yeah. For instance, if you have some sort of tree or class analysis then that you do really see, OK, we have this mutation and by doing that we can then reach sort of part of the code or not.

What impact does mutation have on the coverage?
This would be it I think. So I kept the second thing deliberately abstract.

Execution

P5.20 AT. What initial seeds triggered a certain behavior? E.g., a bug or crash.

Participant
So in terms of the initial seeds?
The thing that we mentioned there that if we perform an execution it might, of course then tell us something about what parts of the code are executed? and also what kind of seeds we would like to have is there. If we know that some of the inputs definitely triggered some sort of bug. would be helpful to highlight them.

So. What kind of seeds, for instance triggered a certain behavior. For instance, a bug crash. This would be something one could highlight.

P5.21 AT. Does an execution of a test case produce a shift in the search strategy?

P5.22 AT. Does an execution of a test case produce a shift in assignment of energies?

In terms of search strategy, we have the execution and it then tells us what.
Yeah, we have also here the link so we can can keep it rather abstract because the same applies also to the search strategy we have input from the queue now, now it's not the initial seeds necessarily.
But again, if we have certain behavior which is triggered and then it affects the search strategy.
Don't know if we already did this at some other point, but how does an execution impact the produce a shift within the search strategy. So I mean I could write it down.

Participant
Does an execution cause a shift in the search strategy, for instance some sort of stage shift or in terms of parameters that it uses or also possibly the assignment of energies. This is related, so if we have a see that a certain type of execution.

P5.23 AT. Does an execution of a test case produce a shift in mutation types or their extent?

Participant
But yeah, in in fact it is also the same for the mutation. So for instance in terms of mutation types or their extent.

how far reaching are of the invitations?

Where possibly are they applied?