Discussion Part 1: Identifying internal stage compositions

Instrumentation

Participant
The first thing that comes to mind, which is, I'm just brainstorming, right, because it doesn't necessarily has to do really with the internals of fuzzing, but of course something that I learned the hard way when I was working with my own instrumentations was the instrumentation always has effect on the execution.

It's like it's almost to me it feels almost like the pop science way of explaining quantum physics, right? Where observing something has an effect on the observed object like instrumentation has an effect on the execution of the SUT which you want to observe. And if it's only the execution time, you will, yeah, maybe slow it down or. Yeah, do something. So this was the first thing that came to my mind.

Maybe by proxy this will have an effect on the power schedule, right? Because in a way or for some fuzzers the power scheduler schedule is somehow linked to the to some parameters of the execution, right?

Initial Seeds

Participant
There are of course related to mutation in the way that, those seeds are like the starting point from where you begin your mutations. Therefore they will effect on how your mutated inputs will look like, right?

Fuzzers like AFL day often also have this this deterministic stages where they're just bit flip stuff and yeah. And therefore I mean there's an effect on mutation.

Could you just, say again, what the search strategy?

Interviewer
So it uses some heuristics to select interesting seeds earlier.

Participant
Then then, of course, this too.

I would also say power schedule because from my understanding, some power schedules, they look at some metrics of your seeds like you have longer seeds and therefore they will be assigned less power, because then you're fuzzing takes longer and therefore, yeah, if your initial seeds are all very long, then maybe shorter test cases in the process will be regarded better and therefore get more power depending on how the power schedule is implemented actually. But I think there might be effects there.

Interviewer
You say that's it for initial seats, we will move to the next box.

Participant
I mean, yeah, in the end everything is somehow linked to execution, right, I would say, but.
I think initial seeds are very important when it comes to fuzzing and.

Yeah.
There's a lot of range for it's more an intuition than I did any studies on it or something but yet choosing your initial seats will have huge effect on how your fuzzing works.

Interviewer
So do you think initial seats also relates to instrumentation?

Participant
No, I have no, no strong feelings about that.

Search Strategy

Participant
Then of course it has an effect on mutation. It of course has an effect on what will be mutated. It doesn't affect the mutation algorithms themselves, but.

Interviewer
So there's no arrow, right?

Participant
No, no, I won't say so.

I would link it to execution because it will decide what will be executed.

if we decide to take something into the queue it also will have an effect somehow on the power schedule in the way that again these test cases will be assigned to certain energy. And depending on how the schedule is implemented itself, if it's complex enough, maybe this will also influence how other test cases will be evaluated, right? Because sometimes it's there's a comparison which is with some thresholds, like, OK, this is considered a long test case or this is considered an interesting test case and in this will have an effect on how other test cases will be regarded in in the process.

Power Schedule

Participant
This will of course have an effect on the mutations because well, it's for the at least nondeterministic status, all these stages that will depend on how long they will be mutated, or how often. And yeah, therefore I mean this, I think it's pretty obvious that there's link.

Interviewer
I think you said there's there might be a relationship between power schedule and execution. When we were looking at the instrumentation box.

Participant
Yeah, I was just, I wanted to just to say, I'm always tempted to have this relationship

to execution because, In the end, it's all about what will be executed. And yeah, of course there there's a link to execution I would say so too.

Interviewer
Anything on instrumentation, initial seeds and search strategy?

Participant
I would say no.
I mean there's this interrelationship I would say between search strategy and power schedule, but this is mostly reflected already in the link between straight search strategy and power scheduling I would say. It kind of goes both ways a little bit, but yeah, mostly from such strategy to Paul scheduling.

Interviewer
Which for the schedule or are you still thinking about these two boxes?
You did not say anything about our schedule to initial seats or instrumentation.

Participant
For me, those stages are almost like pre stages. Once they are once they are done. They will initially affect the internal workings of the fuzzer. The internal workings of the father won't affect them. Yeah, I would say so.

Mutations

Participant
Let's just get the execution out of the way. I would link that. I think I would always link that.

I think I would also link it to the search strategy and the power schedule. Because, I think the mutation algorithms are well like where the magic happens, right? It's where something new is created and almost like out of thin air and I mean this is the idea of grey-box fuzzing right that you have like a feedback loop in some way on multiple feedback loops. Where you gain information by exploring, use this information again to explore better and the mutation is really where it comes together. I think the both the search strategy and the power schedule, they are, like

the interface for this feedback loop, where OK you do something new and then you look OK what happened and then we will, you know, maybe adapt our search strategy if it's that dynamic or at least evaluate what happened. And so I think it goes for the search strategy and the power schedule. Yeah. That's why I would link them.

Interviewer
OK, I'm assuming you don't have any arrows to it. OK, let's go to the execution.

Execution

Participant
It also it affects the search strategy because after you executed the test case, you will evaluate whether or not it is interesting enough to keep it and mutate it. If something happened, or if you want to, do something with it.

I would say. Same goes for the power schedule.

And then those two indirectly affect the mutation. But I think this is already reflected in the earlier relations we did, there's no direct effect on mutation, from execution. By the evaluation from the search strategy and pause schedule.

Discussion Part 2: Identifying visualization analysis tasks

Instrumentation

P3.1 AT. How would the instrumentation alter the behavior of the execution in comparison to a non-instrumented version of the target?

Participant
One thing is. How would the Instrumentation alter the behavior of the execution in comparison to an un-instrumented version of the of the SUT.

Initial Seeds

P3.2 AT. Is there big difference on how many initial seeds are used and their selection order?

P3.3 AT. How different they are from each other? Will the fuzzer be able to generate them over time?

Participant
Here would be interesting what really is the relation between the selection of the internal seeds and those other things right, like, for example. Is there a big difference on how many Initial seeds I use or by some measurement: how different they are from each other.

P3.4 AT. How impactful are various initial seeds on execution?

Participant
In relation to what will be selected in the process as interesting test cases like for example is there a difference if I have only one initial seed and then if I wait long enough I'll get similar test cases generated by the fuzzer itself through the search strategy and power schedule and mutation algorithm. Or is there a big impact when I use various initial seats and is there some way to measure how different from each other? They have to be to have significant effects on those things?

P3.5 AT. There is a quite a big impact of initial seeds on fuzzing performance - A way to measure this

Participant
So this this goes back to my initial thought that there is quite a big impact from the selections of initial seeds on how the fuzzing will go on internally, right? But I can't really put my finger to it. What will happen? But yeah, I have the feeling that, yeah.

Participant
I think, yeah, I think everything about this would be interesting.

Search Strategy

P3.6 AT. Observable correlation between search strategy and power schedule

Participant
It would be also interesting if there's some observable correlation between search strategy and power schedule the way you select the test cases you want to, for example, put in your queue for further mutation will also affect some values some,

let's say fixed values in the power schedule like what is considered a test case that is worth putting more energy to.

Participant
So if you maybe change your search strategy then with the same test cases they will be assigned different levels of energy if there's if there's a codependency or something like that, that would be interesting.

P3.7 AT. Evaluating fuzzing using classic approaches using code coverage and bugs

Participant
I mean the link to execution mostly I think always somehow goes back to those more classic approaches to evaluating fuzzing, what did happen? Did something interesting happen? Do we have new coverage and stuff like that? This is this will always be interesting, I think.

P3.8 AT. Search strategy based on the end goal (ex., replication for patch validation)

P3.9 AT. How seeds are selected affecting the end goal?

Interviewer
If you just consider search strategy and not the relations now do you? Do you think something is interesting to you?

Participant
From the work I'm doing, It's always interesting to. I mean you can use fuzzers for various tasks, right? The classic way to use a fuzzer is to just uncover some box and potential exploits or security hazards or something like that. But something I also work on is like patch validation where you want to see, okay, I patched something and then hopefully the back. I wanted to fix is now really fixed and then.

Maybe it's not completely fixed in a way, and then you start thinking, OK, maybe we want to use a fuzzing approach to solve this problem and then we have to we have to develop new search strategies because we somehow don't want to maximize coverage or something like that. But we want to replicate something that is similar to the originally fixed bug, but not quite the same. And then fine tune your search

strategy to certain parameters and stuff like that and. Therefore, it's also interesting to be able to have some means of having taking closer look to what's happening there, how things are selected and when if they're selected or not and yeah.

I mean, if you work with fuzzing you always come up with some ideas how things could work and then you implement them and then you just let them run for several hours or days and then you take a look at what happened. And yeah, it's always, a little bit of a mystery. I think this is why you why you targeting this right.

Power Schedule

P3.10 AT. Interesting mutations based on power schedule across different algorithms - How does energy vary across different mutations?

Participant
To me, it's somehow similar to the search strategy in the way that, yeah, the search strategy determines which test cases and the poll schedule, how long or how intense you want to work with them.

It would be interesting to see if you alter your power schedule functional algorithm then if this really helps producing more interesting mutations or not. If it's more or less the same and just you just waste your time by assigning more, more power to something. Because this is what you somehow want to also manage with your power schedule, right you don't want to waste your time with things that yeah won't help. But you also want to give enough attention to potential interesting stuff. And to have some data to see whether or not you came up with a with a good power scheduler or not would be interesting.

Mutations

Participant
This is always really interesting because you come up with some algorithm and say, OK, let's flip some bits. Let's cut the test case in half and put it together backwards and let's see if this was a good idea. And then, you have to feedback loop.

Interviewer

We have just seen an example, right? How different mutation phases work. So that's one of the classic example that can take and similar to that can you think of some questions or interesting things that you want to see?

P3.11 AT. Disabling a certain mutation algorithms - does this lead to same amount of test cases in the queue?
P3.12 AT. Does mutation algorithm relate to the energy assignment in power schedule?
P3.13 AT. How different mutation phases affect the observed interesting behaviors?

Participant

Yeah, for example something like if disabling or enabling a certain mutation algorithm. Let's say you have multiple mutation algorithms and you just disable one of them. If this will lead to the same amount of test cases in the queue for example. Or if this will alter the energy assigned to certain test cases. Maybe because this mutation algorithm doesn't contribute anything to the whole campaign. You don't really know because it's running anyways. And then later you can take a look whether you cover so-and-so many branches or whatever, but you never know if they were covered because, like in your example you showed in the beginning, right? You don't know if they were covered because of, I don't know some deterministic bit flipping where you flip every second bit or two bits or whatever, or if they only were discovered because you put total chaos on your test case and generate some random test case you can't link. I mean you can always just somehow in AFL for example, you can take a look at the queue and that will tell you which stage. But yeah, a closer look would be interesting sometimes.

No, I'm sorry. I'm more or less I'm brainstorming on.

Interviewer

This interview structure is for brainstorming, so you don't need to worry.

Participant

I mean, all those, all those topics and questions they have been in the back of my mind and I was always thinking when I was, working on something like that. Oh, it would be interesting to know why this happened or something like that, but I never put it into a structure like that.

Execution

P3.14 AT. Based on what execution information is the seed selected or energy assigned?

Participant

What caused the decision or the amount of energy the test cases were assigned to.

Interesting things I said earlier when it when I talked about search strategy and power schedule and I linked those two execution, it's just the other way around this time.