From:	Peralta, Rene (Fed)
To:	Calik, Cagdas (IntlAssoc)
Cc:	Peralta, Rene C. (Fed)
Subject:	CA paper review
Date:	Sunday, February 25, 2018 10:30:59 AM

Hi Cagdas,

I submitted the following review. Thanks again for your help.
Rene.

----> review

This looks much better than the original submission. There are a couple of issues that should be clarified or corrected:

(1) Bottom of page 13 : "The initial population is created at random and every experiment is repeated 50 times."

This does not fully describe the distribution of the initial population.
Since we are in the middle of a replication crisis, I think it would be important to describe how the initial population is actually generated.

(2) "In Table 2, we give results for each S-box size. Column Original size gives the size of the target S-box used in the regression, and the other columns give statistics for the obtained results (here, column Min represents the best obtained solution). Remark (sic) that all columns refer to the number of primitives in the GP individuals not multiplied by their implementation weights. Note also that we randomly selected the target input S-boxes among those with the best obtained properties from [30,29]."

The best 4x4 rule from [30] is this one: IF(((v3 NOR v1) XOR v0), v2, v1).
This requires three primitives. Presumably, they are counting all the primitives used in the 4x4 S-box, not just what is necessary for the CA rule. This would yield a size of at most 4x3 = 12 for a 4x4 S-box. However, Table 2 reports an "original size" of 77.

(3) Circuits are not trees, they are DAGS. Hence the statement that the size of CA trees is a good predictor of the area of hardware implementations is somewhat suspect. My guess would have been different. But we are all free to hypothesize, so I am not asking you to change the statement. But can you confirm that indeed you are evolving trees and not circuits or DAGS. I ask because if you are evolving DAGS then the mutation

rules would have to be explained.

Typos:

- page 5 line 37: "... its values of ... "
- equation 10 is missing a parenthesis.
- Theorem 2, I think the domain of the component functions f_i should be n-bit strings, not m-bit strings.

-page 9, line 47 : I think Eq. (16) should be "equations (16) and (17).