Major points for consideration:

- The problem statement is not clear. The objectives are not clear. It is also not clear what the inputs of the problem are and what the expected outputs are. Are we designing a decoder? If so, what do we know of the specifications of the encoder? It is not clear why we are trying to estimate functions of the encoder that one may assume that they are known.
  Taking into account that this is likely to be mostly an issue with presentation more than anything else that can be resolved with a proper introduction to the proposed solution.
- Many explanations are missing, such as the idea of the unique and non unique state sets that are used but not explained. The author explained these concepts in other papers but there are no references to these explanations.
- The approach taken in the solution draws from formal methods but does not appear to be solid, or explanations are missing to show that it is. A few of examples that illustrate this issue:
  - In Equation 1 F_{PC} (p,l,r) is defined as a conjunction of a few terms, one of which is $i \neq i'$, and a conclusion is made that if the conjunction is unsatisfiable then $i = i'$. Logically, there are other ways for this conjunction to fail. It may very well be that they all lead to the same conclusion, but this is not discussed or even mentioned.
  - In Equation 2, on the other hand another conjunction is presented, and this time the conclusion is derived on the case that the conjunction is satisfiable. It is not clear, presented or discussed why the conclusion here is different.
  - Equation 2, the idea of loops is also presented though not discussed. It is not clear why the loops are divided into three different sections or why the definition is such that they need to appear in the same places along the two paths.
  - Algorithm one assumes that p, l and r all have the same values, probably for simplicity and without a significant effect, but the difference from the definitions and equations that do not require this property is not mentioned.

  Such issues make the validity of the arguments very weak. I would suggest that either a full formal approach is taken, with all the laborious work of getting that right, but with the reward of the robustness of the solution, or an engineering approach can be taken that is weaker but easier to adopt. The mix created an illusion of robustness that is not scientific.
- In the experimental results, the authors mention that they are not applying their algorithms to the benchmark of another group because it does not contain control flow mechanism. If I understand this correctly, the simpler problem still falls within the scope of this work. Hence choosing not to use it for experimentation is not justified. At least some comparison to the works of others should be presented.

Minor points for consideration:

- Use of Wikipedia reference is not acceptable. The referencing system is built on accountability for information whereas Wikipedia is the opposite.
- "Determining that the design of the encoder-decoder is one of the most difficult jobs" – it's a big statement. Either provide motivation for it, cite others that explain this, or don't make a big statement.
- Level of explanations is not consistent: on the one hand the authors bother to explain the mining of the $\in$ sign as well as other standard Set theory definitions that are taught at an

undergraduate level. On the other hand some more complicated and relevant matters are not explained at all. Such as the meaning of the states and transitions in Figure 1a, and the meaning of "path 2" visiting states.

- Line 21 on page 2: "this algorithm assumes…" which algorithm is this referring to?
- Line 49 there: "none of the current algorithms" – the specific algorithms under consideration should be listed. It is (in theory) possible that the authors are not aware of ALL current algorithms.
- Lines 50-52 thee: this may be a terminology issue, but there appears to be a contradiction in the description. Either the algorithms assume that something is not possible, or they ensure that it doesn't happen. If they assume then there is no need to ensure.
- What does 'pc' stand for in $F_{PC\ ?}$ (the same question applies to other names used in the paper)
- The perspective of equation 1 is not clear. From a decoder's point of view the o's are known and the i's are not. Yet the equation assumes that we know what i is. Perhaps it should be explained first where the i's are coming from.
- Page 5 line 9 - Equation 1 is not explained well enough, "line 2 is a copy of it" – why is there a copy of it?
- Algorithm 1: perhaps this is a matter of terminology, but it is not clear what the inputs of the algorithm are, or why it is designed in this manner. It appears that the actual inputs to the algorithm should be the o's, or perhaps o's and i's, but not only i.
- Is there an assumption in Algorithm 2 that the inputs were already uniquely identified?
- Section 4.1: the logic is not complete. The fact that R(a,b,0) and R(a,b,1) does not necessarily imply that a and b determine t (and should perhaps be written determine the value of t that satisfies R). R may not be satisfiable for any a and b.
- It is possible that section 4.1 should become an appendix and only use its results for the paper. It is a diversion from the problem on discussion.