Reviewer #1:

===========

There could be several criticisms on how the semantics is defined:

starting from the use of lookahead,

the priority of rewriting rules over reduction rules

> Both of these are needed for it to work as we want it to work.

and to the syntactic constraints on how weak actions appear in the prefix without the ";". Indeed one could re-design the semantics to apply the moving of a weak bond to a strong one directly in the concert rules, without resorting to an extra "pre-congruence".

> I suppose this would be possible, but our system nicely separates the reaction possibility from the state.

The use of synchronisation function gamma is not argued, why you don't use taus? Is there any reason to prefer gammas over taus in this scenario?

*We have added an explanation. In the original design, based on atoms, designing the atoms and the synchronizations separately enables the modelling of the many potential interactions between atoms, and we keep the calculus this way in our new application.*

Also the conclusions are quite meagre. You could give some hints on future work, or on how you can add "rates" to CCB since your ultimate goal is to simulate reactions.

Or you could think to extend the tool in [1] so to account for multiple sites.

> Rates is something I had in mind all the time. I just didn't see much of a use, but it should be easy to do.

> FOr the tools, I think we said we do this if reviewers ask. I coudl do it.

Typos & Comments

Please when introducing the syntax of P, please also add the syntax of \mu.

>That relates to below, we use mu in P, but not in C. Probably use mu for both and define mu?

In figure 3 the condition \* does not have to mention b \in WA

>We can leave that out

Fig 5 \equiv is never defined

>Could we use ::= instead

Definition 2 I m not sure what is \rho

>Indeed missing

Reference [5] is wrong: De Nicola and Montanari did not write the CHAM machine

>I have chanted the authors. I am not sure how this happened, normally I export the citations from the journal. There is a conference publication from the same year (which I used in my phd), not sure which one is better (or both?)

On reversibility and Petri nets there have been newer works this year

>https://www.sciencedirect.com/science/article/pii/S235222082100081X anything else?

l 37 sits --> sites

>Corrected

l 188, when defining contexts you should use \mu.C instead of (s;b).C

>See above

Reviewer #2:

============

The paper gives a fairly detailed account of the modeling of the DNA Mismatch Repair using the Calculus of Covalent Bonding (CCB) developed, in other contexts, by the authors. The motivation for using such calculus is that it allow to undo actions in an out of causal order fashion, namely to reverse an action without having previously undone all the other actions depending on it, and DNA Mismatch Repair apparently needs it. I am using the word "apparently" as it is not clear to me why this problem needs this, or if having this kind of reversibility is a common feature of all (bio)chemical reactions. In fact the introduction is from this point rather unsatisfactory and rather assertive, leaving the impression that everything can be done also in other calculi, which are discussed in the related work section. In this section the calculi appeared in literature are reviewed either because they model (bio)chemical reactions or because they encompass reversibility, and

even after it is not clear why the CCB should be used, namely what is the concrete advantage for using it.

>We can reword that. I guess it was clearer e. g. in my thesis

Now I come to the CCB itself. The bonds are represented by the unique shared key between two actions, but the curious point is that actions may synchronize (aka form a bond) according to a synchronization mapping which is meant only to give an observable, possibly different, action, which vanishes somehow. This choice can be understood if a notion of equivalence is provided, which is not the case in this calculus, even in the paper where it has been introduced. And indeed I do not follow it.

>This is similar to reviewer 1. We need to justify the choice of synchronisation.

Nondeterminism is absent, the only possibility being that two parallel components proceed independently making in this way possible synchronizations available and I find this counterintuitive as if everything would be deterministic then the mismatch to be repaired would not arise, but this is not my main issue here.

*There is actually more non-determinism possible. Firstly, since the actions in a site can happen in any order, their execution is non-deterministic. Secondly, if we have a process involving a weak action, for example like this: (a[1];w).(b).P then there is a non-deterministic choice of breaking a or doing b.*

What I find not properly explained in this paper is why the calculus should be in this way, in some sense I got the impression that rules are tailored for the problem tackled

and they lose the generality that a calculus should have. Now the possibility of introducing a rule tailored to a specific problem is possible as a "shortcut", thus expressible with the other rules, but this does not seem the case here, hence a thorough discussion is needed.

>We should stress that the biological problem was not what we used to design the calculus. That should solve that point.

Some other issues on rules are:

\* In the rule "rev s" it is impossible to have the same key twice (or at least is what I get from the other rules), so what is the purpose of the premise?

>I agree

\* in the rule "concert1" the premise is written in an asymmetric way, namely P -(b)[k]-> P' -a[l]-> P" but Q -[k]-> Q' and Q' -d[l]-> Q". Is there a reason?

>Not as far as I can see (wasn't in my PhD, btw, this must have sneaked in)

\* in the rules "concert2" and "concert3" you write in the premises "U ≡ P | Q" and in the conclusion the U is used for that, so why do not go directly for P | Q avoiding a useless part?

>Would solve reviewer 1's remark as well

\* in the rules "concert2 act" and "concert3 act" you write fsh[k](t) in the premise, but I have not been able to find the "t" in the rules.

>already changed

Another issue that puzzles me is that in the prefix (s) one gets the idea that the order in which the actions are written is irrelevant, but in the end, looking at the rules "move-r" and "move-l" it is not, and I am not able to understand why. An explanation here is needed, unless, but then you have to explain why the position of the weak action in a prefix where no weak action is present after the semicolon is relevant.

*There is actually no assumption of a particular order of the actions. Also the weak action can be in any position. The rules move-l and move-r do not assume a particular position. Their move is not left and right (even though that is the origin of the name) but from weak to strong action and vice versa.*

I want to stress that the main problem I have with the rules is the following one. With respect to the CCB proposed by the authors in the paper "A calculus for local reversibility" and "Local reversibility in a calculus of covalent bonding" there is a kind of proliferation of concert rules, and the examples suggest that they are needed as otherwise the DNA Mismatch Repair could not be modeled. The conclusion is that in the original CCB the specific problem tackled here cannot be modeled, and then the relevant question is which problems can be modeled in this extended CCB (a new name would be a good choice) and I believe that the authors have to justify better what they have done, explaining, for instance, why the problem cannot be modeled using the "old" CCB.

>The extension only is the "more parallel". We can justify this.

Though the paper is well written, I find the necessity of being thoroughly revised, which implies an extensively rewriting, explaining better the following points that I already mentioned:

\* in (bio)chemical reactions bonds are created, and also destroyed. Why has the destruction of a bond to be considered as the reversing of its creation? Reversibility has a precise meaning which is, I believe, a bit different from the one implied by the plain breaking of a bond.

>I could write something here

\* why the calculus should have these rules? The fact that they are used in the examples is not an indication of their usefulness.

>The rules really are in all of our papers.

\* the relation induced by the rules is a stratified one. Wouldn't be better to consider a very abstract one without "labels"?, this would allow to define the synchronization algebra not as a mapping but as a plain relation, which I believe is better.

>?

Please revise the bibliography properly, some papers have titles written differently (pi or the corresponding greek letter as an example) and some others do not have the proper author (e.g. the paper "The Chemical abstract machine" is authored by Berry and Boudol, not by De Nicola and Montanari), check also capital letters.

>Done some and will check once more.

Reviewer #4:

============

- My first concern regards the design decisions for the extensions of the calculus. The paper does not sufficiently motivate these extensions and it is not clear if they have been made (and can be applied) only for the purposes of modeling the specific case study. I also wonder: what are the limitations of the extensions and is it possible to generalize the construct to a form that can fully capture concerted reactions?

>Has it been made for this: Yes. Can it do more: Likely. Should I look for examples: Can it be generalized (to n parties): Should be doable, but I am not sure how to write this in SOS rules.

- Secondly, I have some concerns about the technical aspects of the new rules. First, it is not clear that the rules sufficiently take sites into account. Furthermore, in rule concert3, it appears that while executing a concerted action an action may be reversed in unrelated processes, i.e. processes that were not involved in execution of the weak action. I understand that this was not the intention. I give more details about this in the "Detailed Comments" below.

>Not sure about the sites. For the second part, see below.

- Furthermore, I find the presentation of the paper to have a number of shortcomings. On the one hand, the paper contains a lot of repetition from [25]. In fact most of Section 3 is repeated from [25], including examples pertaining to the original version of the calculus. On the other hand, the new aspects of the calculus are not sufficiently discussed nor motivated and I expect that the paper in its present form will be quite difficult to read for someone with no knowledge of CCB as very little intuitive explanation is given. Regarding this second point, I would recommend to the authors to introduce the calculus and its extension informally through a simple but realistic example in the introduction that will highlight the main idea and modeling approach of the calculus and illustrate the need for the proposed extension.

>We can do that, but if we do not remove other stuff, it gets quite long.

- As far as the case study is concerned, as the authors briefly discuss, important spatial issues are not considered in the model and in fact the model appears to allow transitions that do not take place in reality. It would be good if the paper discusses upfront these limitations and argues about the merits and potential of the approach.

>We could add something.

Detailed comments and typos:

In what follows, \*\* is used for more important points and \* for less important ones.

\* line 8 of Abstract: sits => sites

>Corrected

\*\* In the introduction I would include, as explained before, an example for explaining CCB and motivating the new extension. A clear statement of the paper's contribution is also missing.

> An example for general CCB is not new and would expand the paper. A different example might be possible, but it would make it longer. Perhaps an informal discussion of the example could be there? I still think that this is problematic in the introduction.

\* line 48: is => was

>changed

\* line 63: \pi calculus => \pi-calculus

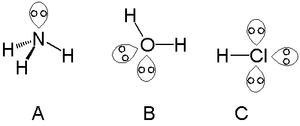
>changed

\* line 112: bee => be

>already corrected

\* lines 147-149: You model a site of a molecule as a sequence of actions that model bonds and a weak bond action. How does this relate to reality? Why just one weak bond action and what exactly is a weak bond action?

This is an explanation of the whole complex, not necessarily for the reviewer. I think we need a sentence or two in the paper like “the concept of weak and strong bonds is taken from electron movements during organic reactions, see our paper [x] for details. Here, we use the same concept without a direct equivalence in reality”. And a bit more for the reviewer. Here’s my long explanation:



The image above shows electron configurations. Notice each bond (line) consists of two electrons. The lone pairs (the bubbles) are two electrons as well. Now there are polarities, i. e. charges distributed unequally in the molecules. If a positively charged part gets close to a lone pair, it can form a bond out of the one pair. Our weak action implements this “attack point”. The atom, which has withdrawn the lone pair, now has one electron “too much”, so it releases another bond (that is the concerted action). Also, the atom where the lone pair came from, is now “missing” electrons, so even if it has a further lone pair, forming of another bond is at least very unlikely (that is why there is only one weak action).

Also, here’s an explanation why examples with two bonds breaking are not very plausible on this level. The new bond is compensated by breaking one bond – that is like for like, so to say.

\* line 166: You mention (s).P. How about \mu.P?

>See reviewer 1, should be uniform.

\* line 170-171: You refer to this feature as "new" but it has already been proposed in [25], etc. The same applies to the reference on line 179.

Reword "The new feature of the operator (s; b).P is the execution of the weak action b, which can happen only after all the actions in s, which are strong actions, have taken place." to "The execution of the weak action b can happen only after all the actions in s, which are strong actions, have taken

place." And change "which is the new feature of our calculus" to "which is the new feature of CCB".

\* line 190: As far as I can see the notion of a subprocess is not used in the sequel.

>Remove

\* line 194: AL is defined to be the set of action labels and then as the forward actions AK, the reverse actions and the concerted actions. However, on lines 128-131 action labels are considered as the set \calA \cup \underline{\calA}.

> Line 128-131 should use a different name. In line 194 ff. I cannot see any contradictions.

Overall, your references to actions and action labels are not the standard ones found in classical process calculi: roughly actions are the objects that decorate transitions. As this can be confusing to the reader, try to be consistent in the names used for the various terms. A table collecting all the notions and notations and how these form the syntax of the language might be useful.

> We might use action labels in 128 and actions in 194 ff.

\* On line 211 in the definition of keys(P) you assume that b could be associated with a key but notation does not permit this.

>It can be temporarily. I believe sos rules allow it.

\* What is the point of the inductive definitions for std and fsn? Couldn't they be defined as:

std(P) iff keys(P) = \emptyset

fsh[m](P) iff n\not\in keys(P)

>keys is inductive as well

\* Rule act1, Figure 3: I don't think you need the premise fsh[k](\mu).

>I think we do to avoid duplicate keys.

\* Rule aux1: what is t in the premise and is this clause really needed?

>should be mu

\*\* Concert rules:

What puzzles me in these rules is that the actions executed e.g. by P and P' need not originate from the same site, e.g. it could be that P = \sigma1 | \sigma2 and the (b)[k] action originates from \sigma1 whereas the \underline(a)[l] action from \sigma2. So, I can't see how you are ensuring that a bond of the same site is reversed.

>Don't really understand this.

By the way, I don't think \equiv has been defined and \alpha is not used according to its intended use as described on line 136. Furthermore, note that the label on R->R' in concert3 does not follow the defined syntax.

>equiv, see reviewer 1 and 2

In rules concert2 and concert3, it appears that if the concerted action involves a weak action in both "molecules" it is sufficient to reverse a strong action in only one of them. Can you give some motivation on that?

>I could motivate that.

And as far as rule concert3 is concerned, I am quite puzzled. It appears to me that its enunciation may not capture the intended meaning. The rule seems to enable the reversal of an action in unrelated processes, i.e. processes that were not involved in execution of the weak action: Q and R could be parallel compositions with distinct subprocesses/sites enabling each of the actions. This may arise since the \alpha transition of Q is not necessarily an auxiliary transition.

>This sound like the comment three above. Do we need to say something about the structure of the processes?

Overall, I find these rules rather ad hoc and not sufficiently discussed or motivated. I also wonder if the rules are sufficient to capture all possible types of concerted actions and, if not, whether they can be elegantly extended to more general concerted rules if it should be needed…

>That is similar to reviewer 2, can we have an n rule?

\* Example 3: There are some typos in the example that could be resolved by setting the second $b$ in the process under study to a $d$ and the $d$ action to $f$.

>There is only one b?

\* concert2+3 act: t should be \eta

>Done

\* concert2+3 res: As far as I understand L does not contain reverse action labels so the side conditions should be corrected.

>This is related to the question of uniform terminology. We say "where L is a set of labels" - we use only "action labels" otherwise, as far as I can see.

\* Figure 8: The last three rules should be upgraded to \mu.P processes.

>Not sure here, we need the structure of mu, so just saying mu can't work, I would say.

\* Move Figures 7 and 8 close to where they are referenced.

>For this, they would need to go before figure 5. This means the SOS rules are no longer in a continuouis sequence. Not sure what to do best here.

\*\* The notions of promotion of actions is not made sufficiently clear. An example of this in terms of a chemical reaction would be useful.

>Again, I can try this, this is a bit similar to others.

Is it possible for any of the strong actions to be used in the promotion? This has confused me in terms of what the actions in s signify? The reduction rules seems to suggest that they are just bonding sites and it is only the keys of the actions that capture bonds.

\* Line 316: the the => the

> Changed (also removed the comma)

\* Line 326: possesSES

>Changed

\*\* It is not clear to me how the properties of reversible computation are inherited from the old CCB to the new one. Some discussion about how the state space is affected by the extensions and new rules, or some other explanation would be useful.

In this respect, I wonder what properties one might prove about concerted transitions.

> I thinks this is all unchanged - any chance to do some proof here?

\* line 354: THE old strand

>Changed

\* page 19: A schematics description of the repair mechanism would be quite useful.

>that would look like our figures. We can try, but it would take half a page or so at least.

\* model description, pages 21 and 22: Please clarify if subscripts are actually needed or just used for easing the presentation/association with graphical representation.

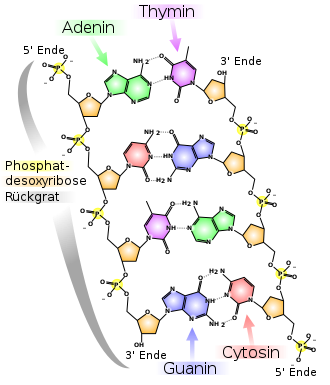
\* line 401: if they ARE parts

>changed

\* page 25, line 7: There is a superfluous S at the beginning of this line

\*\* page 25, last paragraph: How is the correct DP chosen and why is p5 broken as opposed to p3? Similar questions arise for other transitions.

This is explained by combination of the internal structure of the DNA and the spatial configuration. The DNA in more detail looks like this:



Note what we called “DP” is a combination of two groups (P and D, hence our DP) and their combination is not symmetric. The protein UvrD is comparatively large (it probably has a diameter similar to the width of the DNA or so) and its surface is made to match something. So if it is attached to the DNA an a certain position, the site actually interacting with DNA is next to a certain bond (that is the spatial aspect) and it is made to interact with a certain part of the DNA. Only one such part is next to the protein site, and that will be broken.

In a similar fashion, we can explain why e. g. the bond from Guanin to Cytosin is broken and not from Guaning to the backbone – they are not the same (here, one even is a hydrogen bridge) and their spatial location is different.

In general it seems that the model allows transitions that do now occur in reality. It would be good to comment on this and discuss the merits and drawbacks of the model.

\* line 465: extendED

> Changed

Reviewer #5:

============

I found the topic of the paper interesting and relevant. Nonetheless, some aspects of the presentation needs further clarifications.

- Contribution: I think that the presentation needs to stress on the relevance of the contribution. To start with, I would suggest to include a description of the main contributions in the introduction. I wonder whether the emphasis of the contribution is on the proposed extension of CCB or on the case study (as apparently suggested by the current introduction). I would have appreciated a discussion about the shortcomings of the original definition of CCB: does this contribution imply that MMR cannot be modelled with the primitives originally proposed in CCB? Wouldn't be possible to model MMR at the expenses of encoding the behaviour of complex prefixes as low-level (artificial) reactions? I also wonder whether the new primitives are general enough for modelling bio-chemical reactions or ADN problems? In other words, does the new extension of CCB capture the essence of bonding or is it just an ad hoc extension for modelling MMR but may require further extensions to accommodate other phenomena? As far as the modelling of MMR is concern, I wonder whether the proposed formalisation is correct in some sense? It seems that the proposed model can simulate the expected behaviour of MMR, but what about the remaining behaviours (if any)? It is mentioned that the proposed formalisation may allow for additional computations because it misses spatial information. Can we state that additional computations are "only" does that arise from the lack of spatial representation?

>”could MMR not be modelled without the extension?” yes, for my understanding this is correct. Not sure what the artificial reactions mean – I don’t think we can model this “walking with breaking” using existing formalisms. “Is paper about the case study or the extension?” for me, it is about both, they motivate each other. I think we need to decide here and reword the introduction accordingly. “I also wonder whether the new primitives are general enough for modelling bio-chemical reactions or ADN problems” hard to tell, I don’t think this can be said for sure. “Can we state that additional computations are "only" does that arise from the lack of spatial representation?” I would think yes, but proofing this is hard. We could use the software.

- The technical development needs clarifications/adjustments/explanations. In particular,

\* The separation between strong and weak actions, which however can be mixed in the synchronisation algebra (e.g., a weak action may synchronise with a strong one, or two strong actions can be synchronised and become a weak one). What is the rationale? Does this distinction actually match a chemical phenomenon?

>Yes, it does. This should be in our previous papers. I can try to explain it in here, but that would mean more material. A reference may be enough.

\* The non-compositional nature of the semantics. Does this reflect a constraint already present in chemical reactions?

\* The (stratified) semantics for reductions and transitions.

(Detailed comments are provided below)

-------------------

Detailed comments

P1.

L36. Sits -> sites

>Corrected

P5.

L18. "The first attempt at the modelling of undoing of forward computation was RCCS [21]". This is not accurate. I would trace reversibility of computation at least to Bennett, C. H. (1973). Logical reversibility of computation. IBM journal of Research and Development, 17(6), 525-532. ISO 690 which has been reprised afterwards in other papers, e.g., van Tonder lambda calculus.

>This needs rewording. I was referring to CCS, but of course reversible computation is a lot older.

L48. bee

P7.

L23. Fully bonded -> what does it mean? just past actions?

P8.

L12. Definition of gamma. Actions have been divided into strong and weak but gamma allows for mixed synchronisations. Moreover, the synchronisation of two strong actions can give a weak action. This may deserve further explanation. Also, synchronisation would allow multiway synchronisations. Consider gamma defined such that gamma(a,a)= a and P -a[k]-> P', Q -a[k]-> Q' and R -a[k]-> R', then

P -a[k]-> P' Q -a[k]-> Q'

-----------------------------

P | Q -a[k]-> P' | Q' R -a[k]-> R'

----------------------------------------------

P | Q | R -a[k]-> P' | Q' | R

Hence, we have a bond among three components. Is it the case that a bond may connect an arbitrarily large number of components in chemical reactions?

P9

Fig.2. I wonder whether defining sites in prefixes as a commutative monoid with some identity would avoid the "duplication" of rules.

L36. k() -> I would write either k or k(\_) (also for other functions, as keys), otherwise seems to be the application to the unit value (empty tuple).

L41. keys(S) = keys(P) if S = P -> It should be mentioned that recursion should be guarded, otherwise keys will be undefined in some cases. Analogously, for std and fsh.

P10. Fig.3. The communication rule together with the liberal definition of gamma allows for mixed, multiway synchronisations. This may worth an illustrative example.

It is unclear the need of both rules concert1 and concert2. It seems that concert2 subsumes concert1 (after all Q can be taken as a parallel composition of processes). Can you provide an example of a reduction by rule concert2 that cannot be obtained from concert1.

P11. Fig 5. Puzzled about the fourth premise of rule concert3. None of the rules in Fig3-6 allows for the derivation of a transition labelled by a pair of reverse actions; moreover, this label does not belong to the set of groups of concerted actions defined in P8L45 (AK × AK ∪ AK × AK × AK).

Also for the conclusion, it seems that there is a misplaced bracket, i.e., {e[k],f[l]},j[m] instead of {e[k],f[l],j[m]}

It seems to me that there is an underlying design choice that complicates the treatment of reductions (and it is the cause of the proposed extension): The semantics is essentially non-compositional for weak actions. Not being versed on (bio)chemistry, it is unclear to me if this choice actually reflects the essence of biochemical reactions or if it is a shortcoming of the proposed formalisation.

This also concerns the semantics of restriction.

Assume that P | Q | R -{e[k],\underline f[k]}-> is derivable by the application of rule concert2. Now consider (( P | Q)\g | R)\h where:

1- g and h are names different from the ones involved in the derivation of the transition above, i.e., {g,h} \cap {a,b,c,d,e,f} = emptyset where {a,b,c,d} are the names of the actions in the premises of the derivation for the transition above.

2- structural rules do not allow to move out restrictions, i.e., {g,h} \subset (fn(P) \cap fn(Q) \cap (R))

Hence, (( P | Q)\g | R)\h -{e[k],\underline f[k]}-> cannot be derived.

This is quite counter-intuitive since restrictions of names that do not participate in the interaction can block synchronisation. Does this behaviour corresponds to any chemical restriction?

P15.

L9. Concert2 rule makes me wonder about the need of accounting for longer chains. Is it concert2 rule enough or in the future we will need versions involving 4, 5, ... parallel processes? (This is somehow related to the concern about the non-compositional definition of the semantics).

P16.

L43. AK)Proc -> AK) \times Proc

I'm a bit lost about the need of introducing the (stratified?) transition relation of Def.2.

Is Definition 2 saying that reductions have priority over transitions? his is what I understand from Ex 6, but this is hard to grasp from the definition. In such case, wouldn't be easier to state that terms are first normalised (according to =>)? If I correctly understand the proposed formalisation, albeit not confluent, => is strongly normalising. Hence, take nf(P) = {P' | P =>\* P'' =/=>.} and define the (wanted?) transition relation (which I denote with ~~> ) as follows

P ~alpha~> P' if there exists P''\in nf(P) s.t. P'' -alpha-> P'.

Besides what does rho(con(r)) mean?.

Is the provided formalisation relevant for the remaining of this paper?

I would have appreciated a detailed calculation of the Example 6 to understand why Def 2 avoids the unwanted computations.

It seems that -> is overloaded... It may make sense to use a different symbol for the transition relation introduced in Def. 2.

P20.

L48. Definition of gamma. There are cases in which weak actions synchronise with strong actions, e.g., γ(s,p3) = sp3. Is the resulting action strong or weak?

P22.

Terms representing the model without any abstraction are a bit scary and I cannot get why this is meaningful. It seems hard to me to cope with manual verification of the illustrated transitions. Besides, there is not evidence of their derivations. Have they been derived manually?

As far as I understand, the reactions are "local" in the sense that they involve just a few of the components (if we disregard freshness conditions on keys). If this is the case, it may make sense to show only the relevant parts of the derivation of the transition.

P30.

L47-54. If I correctly understand, the proposed model may exhibit more behaviours than those that can actually happen in reality (because of missing information, e.g., spatial information). This makes me wonder about the "correctness" of the modelling. It seems that the proposed representation and the semantics of the calculus suffices to capture (simulates) some expected behaviour and allows for more reductions. The underlying question is: are all expected behaviours allowed in the model? Can we ascertain that extra behaviour arises only because of the missing details about spatial information?

Could spatial information be modelled by recasting to operators in the model? Restriction, renaming of actions, etc?