Identifying Feedback Concepts Using Loops that Matter

By William Schoenberg and Robert Eberlein

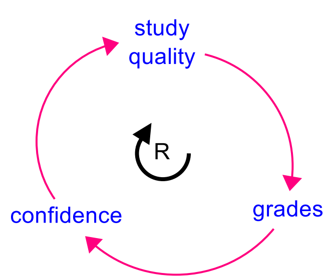
Abstract

When analyzing models using Loops that Matter, there are often two or more loops that basically do the same thing. By combining these loops into a loop concept, and then presenting a smaller number of loops for analysis, the understanding of more complex models becomes much easier. In this paper we show how combining loops based on correlation can streamline model analysis.

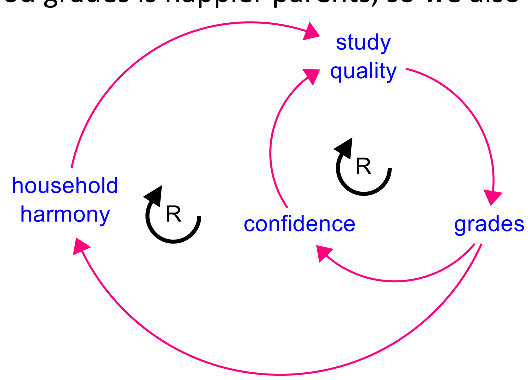
The Problem

The Loops that Matter (LTM) method is used to understand the origins of behavior in system dynamics models (Schoenberg et. al, 2020, Schoenberg et. al, 2021). LTM helps identify the important loops for generating behavior, and thus increases understanding of the relationship between structure and behavior. While this is tremendously helpful for smaller models, where there are a relatively small number of loops active at any time, in bigger models there can be many loops operating where all are significant, but none are clearly dominant. More importantly, many of these loops are likely to be quite similar, both in polarity and in the variables involved.

The propagation of similar loops is a consequence of detail complexity. Consider, for example, the impact of focused study on school performance:



Another consequence of good grades is happier parents, so we also have:



Here we have two very similar loops, both going through study quality and grades. In developing the most compact possible model we might leave one out, recognizing that one takes on the role of both. This makes analysis much easier, but makes it more difficult to explain, and potentially calibrate, the model.

While the above example is simple and lighthearted, if you study models of any complexity, you will see many overlapping loops that largely serve the same dynamic purpose. This is, in fact, something we observed during the development of the Strongest Path Algorithm (SPA) (Eberlein & Schoenberg, 2019). In that case, we were trying to understand which loops get included and which excluded as we restrict the search space for the SPA, and found that they were very similar to loops already included in the sense of both the polarity and the variables involved.

We can think of such similar feedback loops as embodying a single "feedback concept." In our simple example good grades let you be more harmonious with the universe which makes it easier to study leading to better grades. Fortunately, we don't need to name the feedback concept, we just need to be able to determine which loops are working together in their impact of model behavior.

Grouping loops together allows us to combine their importance (as measured using LTM) so that rather than saying there are *n* not very important feedback loops, we can say there is one important feedback concept. That is, fundamentally, what this paper is about.

Let's demonstrate with a relatively simple quantitative example, a three-way arms race model:

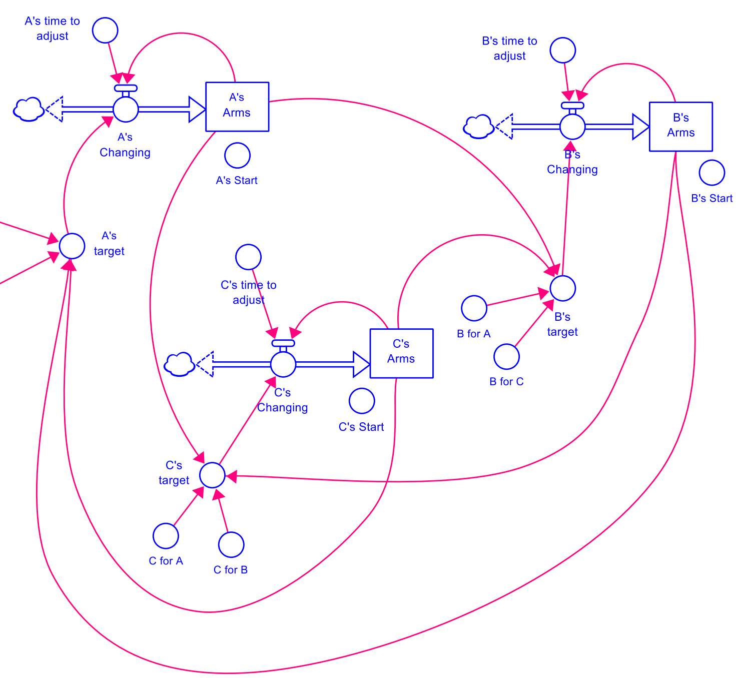


Figure 1: A simple three-party arms race model

The model is set up so that A wants only parity with B and 90% of C, B wants parity with A and 110% of C, and C wants 110% of A and 90% of B. A starts at 50, B at 100 and C at 150. There are 3 balancing stock adjustment loops (the standard balancing loop in the arms race archetype), three pairwise reinforcing loops A, to B's target, to B, to A's target and so on (the standard reinforcing loops in the archetype) and then two reinforcing loops involving all three players (A to B's target to B to C's target to C to A's target and A to C to B (with intermediate)).

When we run an LTM analysis on this model (initialized as described above) we get the following results shown in Figure 2**:**

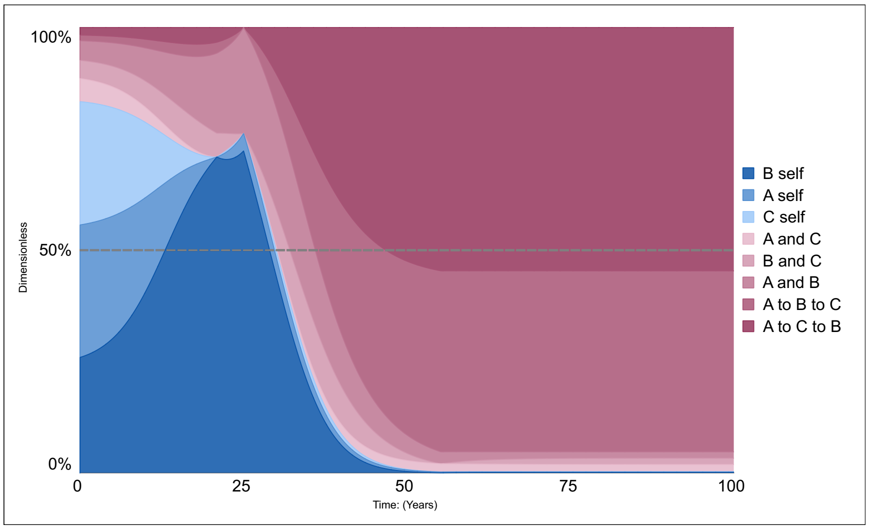


Figure 2: Results of LTM analysis of three way arms race model

If we look closely at the results plotting the two reinforcing loops which link all three players together, we get the plot show in Figure 3:

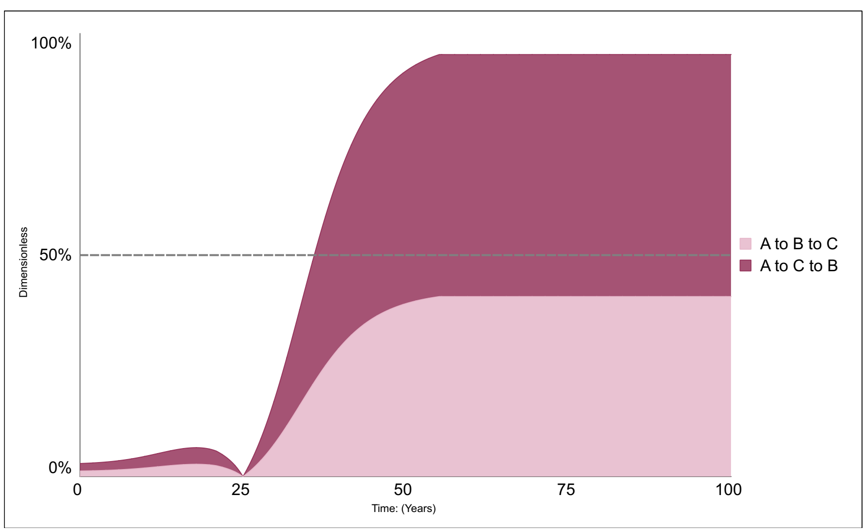


Figure 3: The two reinforcing loops linking all three players

These two reinforcing feedback loops which include all three players represent the same feedback concept. They connect the same variables (in a different sequence), have the same polarity, and are both strong and weak at each point in time. When explaining the origins of behavior in this model, the differences between these two loops are trivial, and instead the behavior of this model after time ~36 is best explained as being dominated by the reinforcing feedback which connects all three players rather than focusing on the differences between these two loops.

When does it make sense to aggregate individual feedback loops into a concept, vs. treating them as individual loops? We have developed an algorithm for that which can be applied to models of any complexity, and dramatically simplifies the loop dominance analysis.

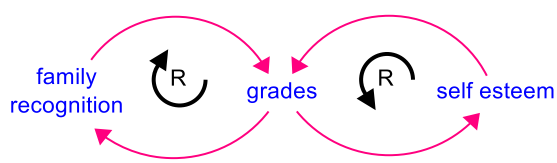
The Solution

Like many things, once the problem has been properly articulated, the solution is clear. In this case, we want to do exactly what is described above, collapse loops that represent similar concepts into single measure. The exact way we do this is constrained by having to maintain simplicity of use of the software, while also allowing enough user customization to work well with a variety of models.

Two loops are considered part of the same loop concept if they:

1. Have the same polarity (throughout the simulation)
2. Share at least one link (are overlapping)
3. Are correlated over the course of the simulation to a user settable threshold

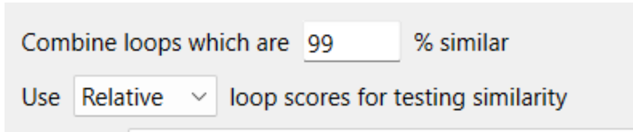
The first of these is pretty obvious. When there are positive and negative feedback loops it is the relative dominance of loops that determines dynamics so we need to keep them separate. The second is sensible but can miss loop concepts that play out through similar, but distinct paths. Consider for example:



This is a variation of the first CLD we presented, and identifies two processes that are conceptually similar, but would not be combined using our rules. The above CLD is also quite abstract, and to create a functioning simulation model from it would likely require adding more concepts, and it is quite likely that doing so would cause the two loops to overlap (as in the first example). For this reason, we are not overly concerned about missing concept matches based on a lack of overlap.

Correlation is a number between 0 and 100 percent (given rule 1) and is easy to measure. The only choice here is whether to use the relative loop scores (those that are shown in the software) or the raw, non-normalized values from which they are derived.

So there are two things to set when grouping loops into concepts:



The first is the correlation threshold, which is 99% by default, and the second whether to use relative or absolute loop scores. Though 99% may seem high, in practice it is quite common for multiple loops to have such a high correlation.

When loops are grouped into concepts, the strongest (as measured by the average loop score over the course of the simulation) is displayed, and the number of loops combined for the concept is also shown, and the strongest loop is chosen to represent the feedback concept and is assigned the combined strength of all its component loops. This decreases the total number of loops requiring analysis making it easier to understand the drivers of model dynamics.

When we run an LTM analysis again on the same three-way arms race model where we combine loops using the default threshold of 99% we get the following results shown in Figure 4, where the two overarching feedback loops are aggregated into a single concept.

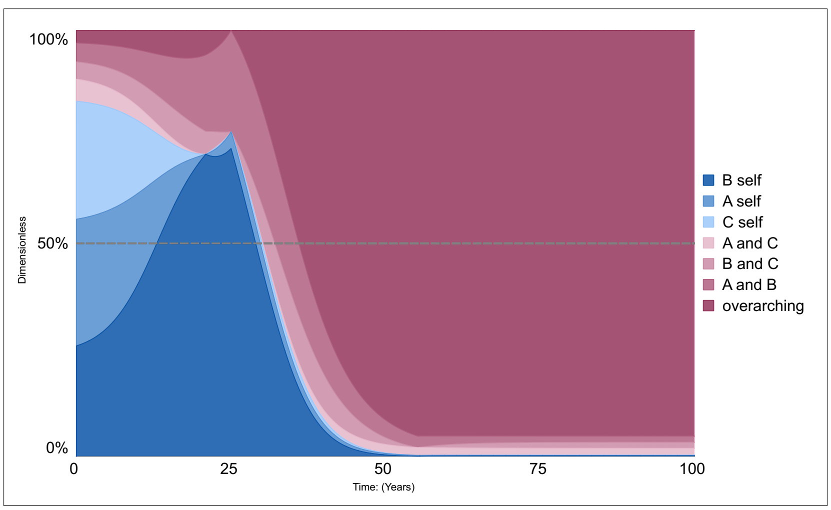


Figure 4: Three way arms race results aggregating the two overarching reinforcing loops into a single feedback concept

Sensitivity to Combination Choices

The first parameter, the correlation level has an obvious impact on the identification of feedback concepts. Any value greater than 100% means that no feedback concepts will ever be identified, and all loops will be treated as individual. At 100%, only loops which are perfectly correlated will be aggregated, and as the correlation level drops more and more loops will be aggregated, and it is therefore up to the analyst to validate that the feedback concepts are relevant and sensical for the model.

The second parameter, the specific metric to use to test for correlation, is far more interesting. Using the relative loop score is far more restrictive when identifying feedback concepts. This is because the non-normalized loop score magnitude for any loop becomes infinity when the second order derivative of any stock in the feedback loop is 0 (Schoenberg et. al, 2021). Because of this, non-normalized loop scores generally are very peaky, and therefore are more easily correlated, because by definition all feedback loops in a feedback loop set share at least one, and typically many stocks in common. Therefore, generally many feedback loops simultaneously are tending towards a non-normalized loop score magnitude of infinity simultaneously which means the correlation approach on these non-normalized loop scores tends to underweight everything but the infinities.

Canceling Loops

The converse of two loops that move together are two loops that have the same strength, but opposite polarity. In this case there is a correlation of -100% between the two loops. As part of the research into combining loops into concepts, we also decided that two such loops, always moving in opposite directions each sharing the same loop score magnitude, are best ignored.

Canceling loops is an option in LTM, one that is on by default. Not all models have them, but those that do become far easier to analyze when the canceling loops are removed. Once they are pulled out, they no longer show up in lists or contribute to the total used in normalizing loop scores. This reduces the amount of information needed to analyze the model and thus increased the ability to understand the model.

Conclusions

System Dynamics models of any significant complexity or detail tend to require the aggregation of multiple trivially different feedback loops into what we've termed 'feedback concepts' to ease the understanding of the loop dominance analysis performed by LTM. The solution we've developed to identify these feedback concepts is to test for high levels of correlation among the relative loop scores of all feedback loops in a model (or feedback loop set). This correlation approach allows the analyst to have control over the feedback aggregation process via the modification of the correlation level parameter, or via changing the metric which the correlation is applied to. All uses of this approach require the analyst to check which feedback loops have been aggregated to validate that true feedback concepts have been identified before the analyst can safely proceed.

References

1. Eberlein, R., & Schoenberg, W. (2020). Finding the Loops that Matter. *arXiv preprint arXiv:2006.08425*.
2. Schoenberg, W., Davidsen, P., & Eberlein, R. (2020). Understanding model behavior using the Loops that Matter method. System Dynamics Review, 36(2), 158-190.
3. Schoenberg, W., Hayward, J., & Eberlein, R. (2021). Improving loops that matter'. In System Dynamics Conference.