A Systems Thinking Primer

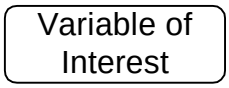
The language of systems thinking consists of just two basic building blocks: the *variable* and the *link*. From these humble pieces we can build quite complex structures. First let's establish a baseline definition of a *system* then proceed from there. Peter Senge wrote:

> A system is a perceived whole whose elements 'hang together' because they continually affect each other over time and operate toward a common purpose.

That's pretty straightforward. That would cover for example the Federal Reserve System, or your cardiovascular system, or an automobile's ignition system. Now to the pieces …

Variables

First, we have the *variable* – by a *variable* we mean simply something whose level or degree or amount can vary over the course of time. That is, from one minute to the next or one month to the next, it could go higher or it could go lower. It can even just stay the same for some period provided that it retains the fundamental characteristic that it *could* go up or down. In grammatical terms, variables tend to fall in the category of noun (Speed) or noun phrase (Actual Volume of Sales). We represent a variable by an enclosed shape such as an rectangle or an oval.



What about 'up' and 'down'? Does this mean that we must have the capability to easily quantify a variable, such as barometric pressure or sales volume? No – *measure* and *quantify* do not mean the same thing. We can evaluate our passion for a certain kind of work – perhaps you formerly really loved your job, but now can barely tolerate it. Has your Morale changed over time? Yes. Can you measure (evaluate) it? Certainly. Can you quantify it? Well, you *could* say, "It used to stand at 9.15 on a one-to-ten scale, but now it stands at only 1.03." But what purpose would that degree of exactitude serve? Can we point to an international standard for morale, like the Official Meter Stick in Paris? Hardly. But for the sake of clarity, you would want to have a clear explanation of just what you mean by the word 'morale.' For the purpose of understanding what *caused* your Morale to change (and in turn what changes come about downstream as a *result* of your reduced Morale), the *direction* of the change takes on clear importance, but the exact degree much less so.[1]

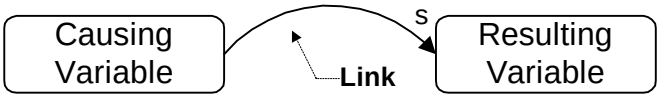
So, a variable exhibits behavior over time in the sense of going up or down, and possesses a highness or lowness at we can measure any given point in time. Note that

[1] This explanation will suffice in the cases where we are dealing with what we might call 'soft' or 'social' systems. If we turn to 'hard' systems, where (a) we can identify with very high confidence all of the relevant variables, (b) we can accurately quantify all of them, and (c) we can determine the weight carried by each of multiple causes, then quantification takes on much more significance. And we cross from the domain of systems *thinking* to the domain of systems *dynamics* and the opportunity to construct sophisticated computer simulations of the model of the system.

some variables we will classify as *internal* to the system, and others we will classify as *external* to the system. More on that in a few moments …

Links

Second, we have the *link*. A *link* represents a cause-and-effect relationship between exactly two variables. That is, we show a link as an arrow: on the blunt end we position the *cause* while on the pointed end we position the *effect*.



Go back to the example about Morale. Suppose that you received a pay cut, and since then your Morale has suffered. We ask this additional question: Could we consider 'pay' a variable, just as we consider Morale a variable? Can your pay vary over time? If it can, then we qualify it as a variable. You might say, then, that a change in your Pay has a link to your change in Morale, it *caused* a change in your Morale. Cause-and-effect, pay-and-morale.

There's more. Not only did the change in your Pay cause the change in your Morale, but you know the *direction* of the change: as your Pay changed (went lower), your Morale did as well. Thus they changed in the *same* direction. We might expect that if your Pay went up, so would your Morale. But maybe you have still more information to capture in your model. Suppose that the Pay has really significant importance to you because you fundamentally feel too much pressure in your job. In fact, they could pay you less if you were subject to less pressure. Ah! Does 'degree of pressure' seem something which can change over time? If we can answer yes, then we have another variable. Here it seems that as Pressure does *down*, then Morale goes *up*: a change in the *opposite* direction.

We need just one more bit of information to finish with the links. While we represent them by an arrow on picture, the diagram, we really need to document just why we believe that we can connect these two variable in this way. We need to answer this kind of question: "Why does it appear that as Pressure goes down, then Morale goes up?" Our answer might go like this: As the Pressure goes up, we start to cut corners on quality and we really want take pride in our work. We would represent this understanding in words like this: "As Pressure goes up then Morale goes down because of the impact on quality and pride in our work." That explains what happens if the Pressure goes up, but what if the Pressure should go down? Our explanation would be backwards. So instead, we could document it this generic way: "As Pressure changes then Morale changes in the opposite direction because of the impact on quality and pride in our work" And when we draw the link, we use the abbreviation 'S' to signal a change in the same direction, and an 'O' to signal a change in the opposite direction.

Mental Models

The words that come after 'because' we call the *mental model* for this particular link. In fact, each link represents a specific mental model of how the system works. We need to

document these mental model as a vital step in constructing the model. And the set of mental models serves as an integral part of the completed systems model. Thus, the systems model consists of both the picture (showing the variable and the links), and the mental models which explain the links. In some cases, where the someone reviewing a variable might not find it self-explanatory, we may also need to include comments on just what we mean by the variables we've used.

We see with the variable called Morale that it represents one effect but results from two separate causes. We could just as easily have the opposite situation – we just as commonly find situations where we have one cause with two effects. In fact, we don't have a hard limit on the number of causes for an effect or vice versa.

Loops – Part One

So far we have looked at what we might call a segment of a large system. We've identified two causes linked to one effect. But we sense that the picture remains incomplete. Morale, after all, must cause something. And something causes Pressure and perhaps even Pay! Without going too deeply into this domain, let's suppose that we discover through our conversation that Pressure results from the number of customer complaints. Hmmm. Can the number of current customer complaints vary over time? Yes. Therefore it qualifies as a variable. And as we understand things so far, when Complaints changes, Pressure changes in the same direction. Now we have a little more complete picture: As Complaints go up, Pressure goes up, then Morale goes down. Now we have the situation where Pressure sits between two other variable. It now serves both an effect and a cause!

Suppose that further conversation focused on what caused the complaints to go up. It turns out that happy staff members provide better service, and better service draws fewer complaints. Aha! Add that to the model by drawing a link from Morale to Complaints. We have a circle, also knows as a *loop* (derived from the idea of feedback loops). As Pressure goes up, Morale goes down; as Morale goes down, Complaints go up; as Complaints go up, Pressure goes up even further, so Morale goes down even further so Complaints go up even higher. A vicious cycle indeed! This kind of loop which feeds on itself we call a *reinforcing loop*. Whichever way it starts it will want to continue. That is, suppose that Pressure were to go down. Then Morale would go up, and the Complaints would go down, so the Pressure would go down even more, and so on. A virtuous cycle! We can summarize the essence like this: A reinforcing loop wants to spiral upward or downward unless something stops it. That something could come some sort of fixed limit (when Complaints get just so high, the company goes bankrupt). Or it could come from a second loop somehow joined to the first one. Only the most elementary (or poorly examined!) systems consist of a single loop. Very few social systems consist of a single loop.

We noted earlier than a single variable may have several links in to it or out of it. That brings us back to the notion of internal and external variables. When we spot something in our system which has arrows in to it or out of it, *but not both*, then we have an *external* variable. That is, it may cause something in our system, or it may result from something in our system, but it does *not* serve as both a cause *and* an effect in our system. Right

now Pay falls into this category as an external variable since it comes into the system as a cause, but nothing in turn seems to cause it.

Loops – Part Two

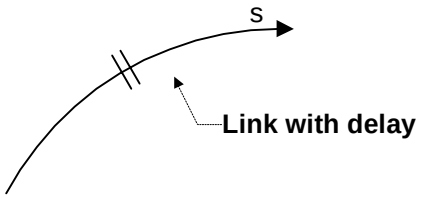
Let's finish up this simple model by adding one more variable. Suppose that during our conversation, we discover that, lo! as the Morale drifts downward, staff members begin to leave the organization. Those few souls who remain now get paid more in order to keep them around. Does number of staff people qualify as a variable? It seems to. Now we have this loop completed: As Morale changes, the Number of Staff People changes in the same direction. As the Number of Staff People changes, the Pay changes in the opposite direction.

Let's walk around this loop and see what story it tells. As Morale goes down, the Number of Staff People goes down because people like to work in an upbeat environment (don't forget to document the mental model!). As the Number of Staff People goes down, the Pay for the remaining staff goes up. As the Pay goes up, the Morale goes up. As the Morale goes up Number of Staff People goes back up again, too. As the Number of Staff People goes up, then the Pay goes down. As the Pay goes down, then the Morale goes down. Note that each time around this loop, the direction of each variable switches over as though there were some kind of magic operating to keep the variables in this loop hovering around a middle value. When they get to some limit, then they switch over. The loop wants to balance things out over time. So we call this loop a *balancing loop*.

Now we have two interconnected loops sharing a common variable, Morale. We might consider them as sort of like paragraphs in a story. When we put all the loops together, we finally get to see the whole story. A good systems model allows us to tell the story. In fact, some people actually prefer to start out by telling the story, then take it apart to uncover the variables the story might hide. We have a reinforcing loop and a balancing loop. If this were the whole story, then we would expect the system to ultimately achieve some sort of balance. We could probably make several trips around the reinforcing loop before the balancing loop kicked in (the people began to flee). We might say, then, that the reinforcing loop has a higher velocity, and tends to dominate over the long haul. That means that the oscillations would have higher amplitude and lower frequency.

Delays

We probably would not see a simple sine wave over a longer period, however. Something else which we have left out of out simple system model causes something more like a 'stairway' kind of phenomenon, like the Dow Jones Industrial Average – some little *ups* and little *downs*, but within a longer trend of *up* (I hope!). Perhaps the Pay goes up, but it takes a while. And when it does finally go up, we have lost track of why it went up (because of the small Number of Staff People). In the case where we think we have identified a cause-and-effect relationship, but where the passage of time somehow confuses us and lets us lose track, we represent that by drawing the double-line across the link. We call it a *delay* but remember that we don't mean just the passage of time. We want to represent the apparent disconnection between cause-and-effect which the passage of time permits.



The triggering question.

Of course, in the real world we could probably connect everything to everything else. Talk about a tough model to draw! So we set a limit on what we choose to include in our model. The best way to limit what we include and what we exclude comes the question which triggered the examination of this system in the first place. Typically, we take the time to examine something because of a problem. Typically, the problem centers on some single thing of particular importance. In this example, perhaps the fundamental issue centered on Morale. So we could launch the conversation with a question like this: "Why does morale seem to get lower and lower?"

This question has three key components. First, it begins with 'why.' We have as our fundamental goal to understand why this happens. Later we can focus on 'what' to do about it or 'how' to treat it. But first we want to understand why it happens. Second, it identifies the variable of primary concerned, and that becomes the center of the conversation: other things have some interest to the extent that they have an impact on Morale. And third, we identify what Morale does that bothers us. Why focus on this variable? Well, we don't want it to get lower. If it were stable or rising, we would not find that a problem. Now it come about that after a while, we decide that something else, some other variable in the system offers us a richer field for inquiry. Maybe we decide to shift our focus to Complaints: Why has the number of complaints risen so rapidly? As long as we all agree to shift the field, that's fine. We almost certainly will wind up with a different model because we have intentionally shifted our question. And the nature of the question determines the nature of the answer.

Wholes and Parts

When we look at complex systems, we see that they frequently form parts of even larger systems. That is, the Federal Reserve System exists as part of the International Monetary System. Your cardiovascular system exists as part of … well, I don't do anatomy. But you get the idea. Because the virtually all of the models I deal with concern *people* we can fruitfully take the definition offered by Peter Senge and add this comment by Russell Ackoff:

> Social systems are systems that have purposes of their own, are made up of parts that have purposes of their own, and are parts of larger systems that also have purposes of their own, and these larger containing systems include other systems that have purposes of their own.

Ackoff makes two key points. Firstly, social systems include parts (e.g., people) who have purposes of their own. Secondly, we might look at the division as the system boundary, or at the department as the system boundary. If we choose the division as the system, then the company becomes the next larger system, the containing system. We call that the *supersystem*. And the department becomes the next smaller system, the

contained system. We call that the *subsystem*. Note that both the super- and subsystem exist as systems in their own rights. Sometimes to help create a context for the system model you really want to concentrate on, you need to create either a super- or a subsystem model to help establish the 'vertical' limits of your model.

You might consider a line of dominoes standing on end as operating like a 'mechanical' system. When you push the first one, the adjacent one topples and so it goes right on down the line. Cause and effect. But if it were *people* standing in line rather than dominoes, at some point one of the people is likely to notice what's happening and shout, 'Look what's happening! Run for your lives!' That's one implication of social systems including parts with purposes of their own.