

The GenomeWeb Intelligence Network



Plectix Debuts 'Rules-Based' Web Platform, Language to Capture Cell Signaling Networks

December 19, 2008 *By Vivien Marx*

Plectix BioSystems, a start-up based in Somerville, Mass., this week launched its first product, a web-based platform and a new formal language system to help academic and drug-discovery researchers represent pathways and study cell signaling systems.

The privately held firm, founded in 2004 by Harvard Medical School systems biology professor Walter Fontana, debuted the software, called Cellucidate, at this week's American Society of Cell Biology Meeting in San Francisco. Plectix currently has 10 employees and a network of six "external academic consultants."

Plectix CEO Paul Edwards explained to *BioInform* via e-mail that Cellucidate "provides access to a rich set of data, collections of data and models relevant to cell signaling systems, coupled with a range of computational tools that enable researchers to appropriately exploit the data."

Cellucidate uses a language that Plectix developed called Kappa that represents molecular interactions and transformations and "makes it easy to assemble or explore functional, virtual representations of complex cell systems even for users without any previous knowledge of Kappa," he said.

In a departure from modeling tools that rely on differential-equation solvers, Plectix uses a "rules-based" approach to address the combinatorial explosion that occurs in modeling biological systems, Edwards said. For example, he noted, modeling the epidermal growth

factor-receptor system requires 10^{33} differential equations, “which is not practical to create or run. Consequently, traditional approaches force the modeler to arbitrarily restrict the activity” that a model explores.

He said his company’s product is unlike those of pathway informatics vendors such as GeneGo and Ingenuity, which focus “on extracting data from published literature relating to genomic and molecular relationships, curating it into maps of signaling pathways, and establishing searchable databases of the interrelations between components.”

This approach, he said, “is useful for orienting research; however it lacks the specificity, syntactic discipline, and the capability of dynamic execution that is required to start investigating the causal underpinnings of disease.”

Playing by the Rules

Harvard’s Fontana explained that the Plectix system uses “rules” to describe biological knowledge. “A rule has a double life,” he said. “It is a description of an empirical finding, [but] at the same time it is an element of a model ... it is executable.”

Rules written in Kappa create a system that can be run like a piece of software. These are “facts with a run button,” he said, which researchers can play out.

When researchers uncover a new fact, for example say, about how two proteins interact, “you can deposit that fact into this large pool of rules,” he said. That not only adds value to the fact, he said, because its consequences run in the context of all the other rules already there, but the pool itself has gained in value because of the added fact.

“It is a network effect where facts in isolation aren’t as helpful as when facts are being put together and being analyzed and their dynamical consequences are being explored together,” he said. “The rules [in Kappa] are like people who join a social network, [but] in this case the rules are facts that join a social network of other facts.”

Scientists building models with these rules can generate a “network of models,” he said, which connects researchers with one another. “It becomes possible, in a system like ours, to make really large models in a cumulative fashion, of complicated signaling systems.” This is akin to large-scale physics experiments, he said, in which the “science transcends the capabilities of an individual.”

The advantage to having a web-based services model rather than downloadable software is that “you can constantly innovate and update,” said Fontana. “You can monitor what the customer wants and as a company adapt very rapidly to the needs of your user base that you couldn’t if the user downloads the software and does something on a laptop [with it].”

Test Run

John Albeck, a computational biologist and post-doctoral researcher at Harvard Medical School, is a Cellucidate beta user. Albeck is using Cellucidate and other tools to look at cell signaling in cancer to possibly tease out factors that determine the transformation of cells into cancer cells. He is using Akt, a single kinase, to test Cellucidate, with the goal of building out the cellular signaling network to involve surrounding pathways such as PI3, phosphoinositide 3-kinase, and mTor, mammalian target of rapamycin.

Albeck and his colleagues have built a Kappa model of the process that activates Akt, and are currently building the corresponding differential-equation model. “Eventually we will decide what set of data is the training set, then we will use that to tweak the parameters of the model,” he said.

Among the next steps is a look at the model to “see if it can make interesting predictions that we wouldn’t have expected just by thinking about it and then go and test those experimentally,” said Albeck.

Albeck previously used Matlab for modeling biological systems, for example, in a study on caspase activation in the cell death pathway published earlier this month in *PLoS Biology*.

The model he developed, called extrinsic apoptosis reaction model, or EARM, includes the species, biochemical reactions, and equations.

Albeck said that he decided to evaluate Cellucidate because he’s looking for a range of modeling tools, and that it appears that Cellucidate might be able to handle larger networks than EARM.

Combine it All

As an example of how Cellucidate might help Albeck in his research, he noted that a kinase may have four phosphorylation sites with activity that

depends on all of them being phosphorylated. But it may bind to a protein irrespective of which site is phosphorylated.

In a model, he said, one would need to tabulate every possibility of each phosphorylation site and all combinations thereof. “If you’re going to write differential equations ... and enumerate all the reactions a molecule could participate in, you have to start writing lots and lots of equations,” he said. “It can take thousands, millions, or even billions of equations to start to enumerate all the possibilities.”

Although he didn’t run into that problem with the caspase network, larger networks risk a combinatorial explosion, which some modeling languages such as Kappa and Los Alamos National Laboratory’s [BioNetGen](#) address, he said. “The way they do it is rather than specifying reactions, which you do with differential equations, you specify rules,” Albeck said.

The difference between BioNetGen and Kappa is “not that big at the linguistic level,” Harvard’s Fontana told *BioInform*. The Los Alamos group has been “pioneering the rule-based approach to modeling,” he said. “Where we differ is [in] what is being done with the language.”

When modeling cellular-signaling systems at the molecular level, the processes that underpin how cells intercept and process signals in response to their environment is “combinatorially complex,” Fontana said. In processes like post-translational protein modifications, cellular components are not “changing identity” but rather taking on different states such as activation, he said.

Cellucidate does not capture how the proteins themselves change on an atomic level, but instead identifies rules that govern the processes themselves. “We are removed from that level of description,” Fontana said. Kappa lets scientists write, for example, that one protein modifies the state of another, emphasizing the network level of signal transmission.

“If we had to describe the stopping of cars at an intersection in front of a red traffic light by taking into account the color of each car, the weight of each car, the number of passengers in each car, then we would really be in bad shape with regard to understanding traffic,” Fontana said.

Traditional tools for dynamically representing and reasoning about networks “break essentially,” he said. Many systems address this challenge by simplifying the complexity of the system, but “we don’t know if we are simplifying in the right way.”

Capturing complexity allows scientists to better understand the principles at work in signaling, he said.

Skip the Equations

For BioNetGen, Los Alamos uses a language to represent the data in which the computer transforms input to generate the underlying model equations. With Kappa, Fontana said he sought to avoid the generation of equations in order to save time and allow for the possibility to reason about the model rather than just postpone the computational challenges of combinatorial complexity.

“Rather than writing these equations in all their detail, why don’t we think of the collection of rules as the model and then take it from there?” he said.

A rule, said Albeck, looks like a reaction with the left-hand side of reactants and a right-hand side with products, but it is “context-independent.” With Kappa, researchers can say, for example, “make this reaction occur, but I don’t care about what combination of phospho sites [are] on the kinase,” he said. The software will then model the reaction independent of the phosphorylation sites, he added.

“A differential-equation solver won’t know that is what you mean; it will only do exactly what is written,” he said, whereas Kappa allows for “shorthand.”

Scientists can also extend beyond the shorthand in ways equations do not allow. “You can say, ‘I don’t care about these four sites, except for this one case when these three are phosphorylated and this one is not’ ... so it cuts out a lot of the work in terms of specifying the full network and all the possible reactions that can occur,” he said.

Each Kappa model more or less represents a single cell and models can run with deterministic or stochastic simulations, he said.

“With Kappa, at least in the Kappa factory program that Plectix has built, the default is to run a stochastic simulation, which means that every time you run it you might get something different,” Albeck said, as the solver is “sort of replicating what happens in the cell” as it runs through different possible combinatorial possibilities.

Models don’t replace the wet lab and experimental validation, but they help to “look at things in a moving parts kind of way,” Albeck said. Much like

looking at the parts of an airplane engine, “you don’t quite understand how the engine works because there is a process going through the entire engine that everything is a part of.”

Running models through Cellucidate can help researchers get a sense how a pathway may or may not run, and allow them to run through the possibilities relatively quickly, Albeck said. “By doing it a little more formally rather than informally like we used to [do modeling], it has the advantage of more power and you can have more faith that your thought experiments actually mean something.”

© Copyright 2008 GenomeWeb Daily News. All rights Reserved.