Research announcement:

PLASMA
The Princeton Locomotive and Shop Management System

After over 12 years of research, we are ready to announce the development of PLASMA, a new planning system for locomotives which has been under development at CASTLE Lab at Princeton University. The system has been under development for 12 years, primarily with support from Norfolk Southern but also with four years of support from BNSF. The system represents a major breakthrough in modeling and algorithmic technology, developed specifically with the needs of large-scale freight transportation in mind. NS has issued the following statement:

“PLASMA has recently completed the user acceptance test at Norfolk Southern as a strategic planning system. It has been used to assist the decision making in the 2009 locomotive road fleet requirement. Reaching this critical plateau required producing a carefully calibrated system that captured NS operating policies while matching their performance on important measures such as train delay.”

The technology, known as approximate dynamic programming, combines the power of linear programming with the flexibility of simulation in an integrated, coherent framework. It is an optimization technology, but unlike the older “optimization models” (based on linear programming), approximate dynamic programming is a technology that captures not just real-world operations at a high level of detail. In addition, it has also been found to capture the intelligence of real-world decision making. In this way, it is often viewed as a simulation technology. We often refer to it as an optimizing-simulator.

PLASMA distinguishes itself from older modeling technologies in the following ways:

- It is able to model locomotives and trains at a very high level of detail. We simply have not encountered an issue that we cannot handle.
- PLASMA can be run on deterministic datasets, or it can be used to model different forms of uncertainty such as transit times, equipment failures, changes in train tonnages, and last-minute schedule changes. PLASMA uses robust optimization technology to design power management strategies that handle uncertainty.
Despite the large scale of the locomotive optimization problem, PLASMA is able to use commercial optimization solvers such as Cplex to produce consistent, robust solutions.

PLASMA is based on an algorithmic technology known as approximate dynamic programming, which has evolved in the artificial intelligence community as a way of mimicking human behavior, and within the operations research community as a way of solving optimization problems under uncertainty. PLASMA brings these two fields together to produce realistic, intelligence solutions that behave the way real operations behave.

PLASMA has three applications:

1. Strategic planning, where a highly-detailed model of operations is needed to study a wide range of business issues ranging from fleet size and mix, management of foreign power, policies regarding consist management, maintenance policies, and the study of transit time and equipment reliability.

2. Short-term operational planning, producing forecasts of locomotive and train movements over a 1-5 day horizon to help identify and fix locomotive shortages, and assess the impact of last-minute schedule changes.

3. Real-time locomotive optimization. While this use is still in the planning phase, this application takes advantage of our ability to re-optimize locomotive assignments in a few seconds as users work with the system interactively.

At this time, we are focusing primarily on strategic applications. Questions that the strategic version of PLASMA might be used to solve include:

- What is the correct fleet size and mix for to serve a projected train schedule in the future, while holding the level of train delay constant?
- What is the value of changes in fleet mix on network performance?
- What is the impact of reductions in transit time variability on service performance and fleet size?
- What is the impact of changing the location or capacity of maintenance shops?
- What is the effect of changes in policies for handling foreign power?

Although our interest is in strategic planning applications at this time, we note that PLASMA is easily adapted to run operationally, providing tactical forecasts 1-7 days into the future with updates every few minutes. A version of the system is also planned which can run interactively for real-time locomotive planning.

The history

The development of PLASMA has its roots in other projects as we found that traditional optimization models simply were not working for the types of large-scale applications that arise in freight transportation. The algorithmic technology has evolved considerably from the early 1990’s when it was first applied to model the flows of drivers for major
trucking companies. The project to apply these ideas to locomotives began in 1996 with Norfolk Southern, and in 1998 with a parallel project at BNSF. The project at BNSF ended up focusing on shop routing, which allowed us to develop considerably expertise both in the modeling of western railroads, as well as the challenges of developing a truly powerful shop routing system. However, the development at NS stopped in 2002 due to various reasons.

During the 2003-2005 period, we made several major changes to the underlying library which have dramatically improved the stability of the system and solution quality. Starting in 2006, we were given an opportunity to implement this new system, now called PLASMA, at NS in a project with three phases: 1) strategic planning, primarily for fleet sizing, 2) short-term tactical planning and forecasting, and 3) real-time locomotive optimization. As of this writing, the fleet sizing system has been implemented and accepted by management.

The current project, started in 2006, has been conducted with the close collaboration of the operations research group at Norfolk Southern. The locomotive planning system at NS is called LARS (Locomotive Assignment and Routing System), which includes PLASMA along with an extensive set of data collection and reporting capabilities.

The project at BNSF focused purely on real-time locomotive optimization. We were able to meet or exceed every operating performance measure, including some particularly stringent measures on shop-routing performance. However, the IT integration never came close to meeting the needs of a real-time system. At the moment, we are focusing on the use of PLASMA for strategic and short-term tactical planning applications.

Capabilities and features

PLASMA is able to model operations at an exceptionally high level of detail. We have found that even for strategic planning, a properly calibrated model requires an accurate representation of locomotive operations.

Some features of PLASMA include:

- Each locomotive is modeled individually, making it possible to capture accurate horsepower ratings, locomotive features such as cab signal, LSL, flush toilets and even bullet-proof glass.
- Arrival times of locomotives, departure times of trains and set-off times can be modeled to the minute.
- Consist breaking is modeled properly, and penalties can be adjusted to keep the rate at which locomotives are set out or swapped to historical levels.
- Each train is characterized by minimum horsepower requirements, goal horsepower requirements (to meet service) and a maximum on the number of locomotives that can move a train.
- Trains are characterized by precise tonnage requirements and service expectations. We can enforce a minimum horsepower requirement to move the train, and a goal horsepower to make service. The model will balance the costs of
breaking a consist with the benefits of spreading power around so that every train meets its minimum and as many trains as possible hit their goal horsepower.

- The management of foreign power can be modeled to closely match railroad policies. The model can be adjusted to encourage or discourage the use of foreign power depending on whether the railroad is in a surplus or deficit situation.

- Locomotive repositioning can be accomplished either by adding extra power to a train, or through “light engine moves” where locomotives are moved by themselves. The user has complete control over where light engine moves are used, and when. It is also possible to impose limits on the number of locomotives used for a light-engine move.

- The model has sophisticated shop management capabilities, primarily as a result of two years of development with BNSF. We can guide power to a home shop (or a shop for a specific maintenance requirement), while simultaneously using locomotives moving toward shop in a productive way. The system will simultaneously estimate delays at shop, and can use estimated delays to route power toward less congested shops.

**Recent applications**

We have been using ADP for several other projects. In a recent project at Schneider National, we used ADP to simulate the behavior of Schneider National’s fleet of drivers. A paper describing this project is available on the front page of the CASTLE website at [www.castlelab.princeton.edu](http://www.castlelab.princeton.edu), or by clicking here).

In a separate project, ADP has been successfully deployed at a major manufacturer of regional aircraft to manage inventories of their spare parts.

We have also used ADP for numerous studies in military logistics, including an analysis of the value of robust optimization. A copy of papers describing these projects is available on request from powell@princeton.edu.

**The technology**

What makes PLASMA unique is the way that it integrates classical optimization and classical simulation in an integrated way within a framework known as *approximate dynamic programming* (ADP). ADP has its roots in the 1950’s, but has evolved independently in separate communities. Within operations research, ADP has been viewed primarily as a method for solving difficult stochastic optimization problems. But the artificial intelligence community has been using it (also since the 1950’s) as a technology for capturing how people solve complex problems. Our own introduction to ADP arose from a need to decompose very-large scale (but deterministic) optimization problems arising in freight transportation. A thorough introduction to approximate dynamic programming, largely influenced by these freight transportation projects, can be found in *Approximate Dynamic Programming: Solving the curses of dimensionality* (John Wiley and Sons, 2007). See [www.castlelab.princeton.edu/adp.htm](http://www.castlelab.princeton.edu/adp.htm) for more information.

© 2008 Trustees of Princeton University
PLASMA uses iterative learning to estimate the value of different types of locomotives in the future. This learning is captured through a device known as a value function approximation which estimates the value of locomotives in the future. It is through these value functions that PLASMA can produce behaviors that work well over time. The value functions also keep individual problems small enough for commercial solvers, and also make it very easy to handle uncertainty.

ADP is more than just an algorithm; it opens the door to a much richer way of modeling problems than was previously possible using classical optimization techniques. Instead of solving a single large problem, it breaks problems into pieces that are solved over time. The logic steps forward in time just as a simulator would, making it possible to draw on the flexibility of classical simulation. For this reason, it is useful to think of ADP as an “optimizing simulator.” It is this technology that allows us to model operations at such a high level of detail.

ADP solves problems much the way people do, but within a mathematically rigorous framework. The technology is very intuitive, reflecting its origins as a method for replicating human decision making. But we have also learned that people are very smart, and it is necessary to have a very smart technology if we are going to even replicate human decisions. It is a common misperception that people make a lot of mistakes in operations. While mistakes are certainly made, we have generally found that people are very resourceful, and we have needed all the power of our optimization tools to match historical performance.

Modeling the organization and flow of information and decisions

PLASMA has a unique ability to model not just the flow of locomotives and trains, but also the actual organization of how decisions are made, and the arrival of information over time in the form of last-minute schedule changes, train delays and equipment failures.

PLASMA models the organization of decisions by breaking the locomotive management problem into a series of subproblems. PLASMA runs in three modes: system mode, desk mode and yard mode.

In system mode, we optimize the assignment of locomotives and trains at a point in time as if there is a single planner running the entire railroad. This produces the best results, but they are harder to implement because this is not how a railroad actually works.

The model can also be run in desk mode, where a “desk” literally refers to a single planner managing a particular region of the network (at his desk). In this mode, the network is broken into regional subproblems. The model is still able to reposition power between regions, just as planners will coordinate their activities, but the degree of coordination is not as high as the system mode. Perhaps not surprisingly, this mode most closely matches actual performance in our work.

The model also captures the flow of information in a realistic way. For example, we can model that a train was added to the schedule with 36 hours notice, or that a cross-country train that normally requires 72 hours transit time was delayed 24 hours into the trip,

© 2008 Trustees of Princeton University
giving us 48 hours advance warning of a delay to incoming locomotives. This is much more than just modeling random events; we model when the information becomes known, which allows us to design appropriate response strategies.

The model can be run in both deterministic mode (where we assume that everything is known in advance) and stochastic mode (where we explicitly model uncertainty). In stochastic mode, we can capture uncertainty in the train schedule (additions and deletions), changes in train tonnages, transit delays and equipment failures. We can also model uncertainty in the arrival of own and foreign power from other railroads.

When run in stochastic mode, the model will exhibit robust decision-making behavior. For example, we may provide 14 locomotives to a yard that might only need 10 or 12 locomotives given the anticipated train departures over a day. But the possibility of adding trains, or late arrivals of inbound trains, may mean that additional power may be needed. Without robust behavior, random events can produce complete schedule breakdowns, something that would not happen in actual operations simply because planners take similar precautions.

The software
PLASMA has been written using the Java-based DRMS modeling library developed at CASTLE Laboratory at Princeton University. The DRMS modeling library is designed for maintenance, including in-house maintenance at a railroad. The library is also well suited to be adapted to other applications, including freight car management (also running at NS), intermodal applications (our first application in railroads in the early 1990’s), and crews. We have separately applied the library successfully to major projects at Schneider National, Netjets and Embraer.

A powerful dimension of PLASMA is its ability to use a core piece of software for assigning locomotives to trains for strategic planning, short-term (1-5 days) operational forecasting and real-time locomotive planning. If the system is implemented as an operational tool, improvements made under the scrutiny of real-time planning transfer immediately to the strategic planning system.

PLASMA is combined with PILOTVIEW, a powerful graphical diagnostic system that allows analysts to see the insides of the model at a high level of detail. PILOTVIEW makes it possible to see not just the decisions that PLASMA made, but the decisions that it considered but did not make. Users can do drill-downs to see the attributes of not just the locomotives and trains, but also information on each possible assignment of a locomotive to a train. We have found that experienced locomotive planners always have a sense of what the solution should look like, and they often are asking “why didn’t the model do this.” We make it possible to perform detailed analyses of the “decision not made” to understand why the model behaved as it did.
Next steps

PLASMA is ready to be tested at other railroads, although at this stage our primary interest is to use it for planning studies. We have learned through our different projects that each railroad is unique, and as a result we do not anticipate that any locomotive optimization model can be used as an off-the-shelf system. However, we do feel that we can do an initial demonstration very quickly that will make it possible to assess the amount of work required to calibrate the model to an individual railroad. The level of calibration also depends very much on the nature of the planning question.

If you are interested in learning more about PLASMA, contact me at powell@princeton.edu to set up a time where we can discuss the system.

Warren B. Powell
Director, CASTLE Laboratory
Princeton University