Extending the RMS Paradigm of Model-Based Computing to Meet Requirements and Demands of Real-Time, Real-World Emergent Critical Processes

Martin J. Dudziak, PhD
TETRAD Technologies Group, Inc.
14.June.2007

Early DRAFT version

Introduction

MPC and RMS

Model-based computing (MPC) is hardly a new concept in computing and arguably has its foundations in some of the early generations of artificial intelligence research going back to the 1970's and 1980's []. Qualitative physics [] and medical expert systems (e.g., Mycin []) are but two examples of model-based computing from an era of single-core microprocessors, LISP machines and RETI algorithms. One of the severe limitations of early AI and subsequent neural network and other statistics-based learning systems was the problem of tackling very large data sets in search and recognition algorithms; the need for trade-offs to accommodate processing limitations (including both processor speed and memory) brought a steep cost in meeting real-time demands or in reducing either false positives, false negatives, or both. The availability and affordability of computing resources that enable supercomputing capabilities to be applied to a large number of business and consumer applications now enables realistic MPC to be considered as more than an exercise for computer science researchers.

One of the more well-known representatives of MPC today (2007) is in the Recognition-Mining-Synthesis (RMS) paradigm advanced principally by Intel. RMS is an application software model for terascale computing and in particular the subject of research and development within Intel for its multi-core processor architecture []. The Intel Tera-scale architecture itself manifests a course of development in parallelism that can be traced to MIMD (multiple-instruction-multiple-data) platforms in early parallel processing that originated with the Inmos T-series transputer (T414, T800, T9000) and formally, even early, in the CSP (communicating sequential processing) models of C. A. R. Hoare and others [] in the early 1980s. Following a decade-plus love affair with shrinking photolithographic technology and increased processor speeds, the semiconductor industry began to address energy and performance efficiency by building from the foundations anew, with a change in scale such that parallelism, workload distribution, fault-tolerance (spares), local specific dedicated functions, and interconnect fabrics are now being down on-chip, on-die, rather than on-board. What goes around - gets smaller - then comes around all the same.

Today RMS shows great promise as the progenitor of a class of MPC architectures that can address a multiplicity of problems that have existed for a long time, if not forever, in the world, but which are dramatically demanding solutions which can only be assisted by high performance computing on a massively distributed and commonplace scale; i.e., within the computing world of homes, automobiles and public environments. This class of problem is one that is directly linked with the confluence of several factors and situations, ranging from the high human population on the planet and the increased affluence (esp. mechanization and mobility) of that population; energy demands and global climate change including a rise in nonlinear weather phenomena including major storms; an increase in the average lifespan, an increase in environmental hazards and pollutants, and a consequence increase in a variety of challenging; and a dramatic increase in local and global armed conflict including the easy facilitation of terrorism and acts of mass destruction. There are problems today that are more demanding of solutions that require massive data space searches, pattern matching trials and evaluations, interpolations and fittings among complex n-dimensional objects, and synthetical reasoning, the complement of analysis and invariable a more complex process for computational systems because of the open-ended paths that need to be explored in order to determine some optimal conclusion, whether that be in face recognition, obstacle avoidance, molecular dynamics, or personal video management. Many of these problems are extraordinarily "critical" for the lives of millions, either because the problem is something that can affect large numbers of peoples' lives at once (e.g., situations concerning climate change, pandemics or terrorism), or because many face the same issues (e.g., traffic safety or individual health maintenance). Others are less critical in the sense of life-or-death but are perceived as demands and needs by a society of consumers that has become accustomed to living with increasing masses of personal data such as audio and video, all of which needs the power offered by an RMS model of computing in order to satisfy the individual consumer. The attention in this brief paper, however, will be upon the former class of such critical problems, namely those that can make differences of life or death for large numbers of people in their individual lives or en masse together.

ECP

The term for this class of problems is Emergent Critical Process or ECP (also referred to as Emergent Critical Event or ECE). These are inherently nonlinear and catastrophic from a mathematical perspective, and generally they have the potential of being catastrophic on physical, social, and economic levels when they occur and especially when they occur in manners for which the subject population is unaware and unprepared. ECP problems may be modeled as simple bifurcations and two-dimensional flows, or as s-called strange attractors (e.g., Lorenrz, Rössler, Tamari, Hénon) or more likely in mass-social settings, the model will be itself composite and involving interactions among a set of different behaviors, some of which may be quite linear and others frankly non-algorithmic and requiring good heuristic judgment for approximations. The December, 2004 Indian Ocean Tsunami is one case in point. Almost any earthquake is another. Hurricane Katrina, albeit an ECE with much forewarning and plenty of computer simulations and predictions for its coming, was a classic ECP. Intentionally-triggered events such as the 9-11 WTC attack is yet another type of ECP. In all these cases the ECP is not "simply" the physical event such as a hurricane, tornado, earthquake or explosion that can be modeled in reasonable fashion well enough, but a far larger set of events that involve massive numbers of components such as individual persons, vehicles, and systems which are interacting and with a variety of dependency relationships, not all of which can be expected to be evident early-on in the modeling process even for subject experts. This means that the element of discovery (novelty, innovation) is something that must be considered in many computational approaches to managing ECP, and this is a dimension of the classic AI (or simply "I" for Intelligence) problem that needs to be addressed in thinking about paradigms for model-based computing.

ECP and specific ECE are not, however, limited to massive newsbreaker phenomena that affect thousands or millions of people at once. They may be extremely localized in scale and in type of effect upon people's lives. These are ECP in the home, in the automobile, in the body, in everyday affairs of life. There are many ECP that happen every day and night, and these are the events that provide a very large and open-ended set of application opportunities by which hardware and software developers, and the developers of many consumer products, can and should take advantage of MPC and RMS-type models in particular. These are also the types of ECP computing challenges that, if intelligently addressed for not only processing capabilities and programming methods but also for human-machine interfaces and the proverbial "user friendliness" will open up massive opportunities for high performance computing (HPC) in the general consumer world.

ECP that occur in the home, on the road, in the community, and in the body are the subject of this brief memo wherein are presented some modifications to the RMS paradigm and some suggestions for generalizing MPC. The objective of making some changes and creating an abstract MPC is to create a set of problem-solving methods and tools that can be employed in use-case modeling and code generation for use in a very wide range of applications, for using such processors as the Tera-scale devices engineered by Intel, without forcing developers to continually reinvent basic techniques and code. The complexities of programming an 8-core, 80-core or larger single-die system are severe and demanding for most programmers. The complexities of the applications are even more demanding, especially if one considers that by virtue of the large number of input devices or data streams, data mining load, and synthesis outcomes, there is likely to be a frequent need to add, subtract, and modify processes that must be performed in the RMS type of application. Unless the overall application can easily accommodate modification and retrofitting of new input, processing and output types, the computational values and the economic benefits associated with Tera-scale computing will be significantly cut back by losses in optimal performance and costs in modifications to the application software. Lest the system be cast back into the type of problems that beset mainframe computing and pre-object-oriented software development of the 1970's and 1980's, it will be important to develop an abstract RMS application development tool that can be used easily for building new systems and for modifying existing applications. The paradigm is something like but not exactly like the APIs and database query tools that have evolved in conventional applications software.

Completing RMS → DRMSO

The model of RMS is effective but perhaps a bit incomplete and this needs to be addressed briefly before proceeding. Recognition-Mining-Synthesis is only part of the picture, a very central part, but it omits two important process components that cannot be easily lumped together into this model. Ahead of recognition there is Detection. It can be called, variously, sensing, data collection, observation, noticing, gathering, or triggering. It is the process, potentially very computationally intensive, of finding that-which-is-to-be-recognized, the pattern that is identifiable as a pattern, or as an anomaly, or as a trigger, but which is not yet recognized or classified. Examples will help to clarify this process as well as its relationship with Recognition and its importance for Mining, Synthesis, and the second mostly-omitted process component, Optimization. After Synthesis, there is generally, in most applications, not a final static state, but an iteration or a continuation or the Synthesis process with new information that may be a return to some part of the Detection or Recognition phases. Often there is a need, or at least an appropriate place, for optimizing that which has been synthesized and which is being put into action as a response, an output, a final throughput from the rest of the process pipeline. Again, it is hopes that concrete examples will help in clarification and in understanding relationships.

Four Types of ECP

Four ECP are considered in this initial exploration of abstraction for the same of tool-building, each in a specific area of large-scale applications for computing in 2010 and beyond [[1]].

1. Security – an example from public transportation security and public health
2. Health – an example from cancer diagnostics & treatment (deformable image registration)

[1] The reality is that these ECP exist intensively and broadly right now and have existed for some time. Intel's plans for its Tera-scale processors appears to be in the vicinity of 2010 or soon thereafter for large-scale commercialization. Let us say that 2010 marks the likeliest early date for bringing Tera-scale into the marketplace.

3. Energy – an example from the automotive transportation world
4. Lifestyle (Home) – an example from home energy and safety control

Each of these four examples will be considered in great brevity, since each deserves, and requires, a very in depth analysis and formal study. The purpose of this memo is to lay the groundwork for future work that can accomplish such a study and an implementation using a Tera-scale architecture.

Following a presentation of the four examples, there is a description of one possible formal model for ECP applications using a DRMSO (Detection-Recognition-Mining-Synthesis-Optimization) model, and some suggestions as to how such a model can be implemented as a software tool for application developers. The four examples are summarized thus and are drawn from prior and current work initiated and led by the author:

Security – Nomad Eyes™
Objective – Identify and track a suspected terrorist or an unaware carrier of a contagious disease through a public transportation system and provide information regarding present whereabouts, past itinerary, and past contacts with associates or bystanders.

Health – I^3DIT™
Objective – Process diagnostics and/or therapeutics for potential micro-tumors or other tissue pathology, on the basis of multimodal electromagnetic and acoustic scanning coupled with nanoscalar molecular medicine.

Energy – I-Trans™
Objective – Enhance and optimize the performance and safety of passengers and vehicles in a hybrid rail-highway transportation network for personal automobiles and trucks operating in hybrid fuel and/or multi-engine systems[2].

Lifestyle – "Orbis-Uno" – the Smart Healthy Safe Home
Objective - Optimize energy, security, and comfort systems in a home on the basis of intelligent real-time processing of information about the physical environment, occupants, and activities within the home.

[2] Although this is a futuristic scenario, the application includes, de facto, many of the requirements for handling ECP situations that would arise in passenger automobiles operating on streets and highways in conventional (2007) traffic conditions and driving patterns. This I-Trans model pushes the DRMSO requirements to the max and also introduces a very practical and feasible solution for carbon emission reduction and "green" transportation that, unlike any other known model, does not conflict with the economic interests of the automotive and fuel industries or the psychological drives and dispositions of the general population in many economically developed countries.

Security – Nomad Eyes™ []
This example is drawn from the need to improve real-time intelligence and avoidance procedures for circumventing an ECP that could result in a terrorist attack or a non-terrorist incident of equivalent or greater effect such as the inadvertent spread of a highly infectious disease such as a human-transmissible bird flu or a drug-resistant tuberculosis. The case model is one that can be easily applied to other computational scenarios in smaller settings, such as the security environment for a school, office, hospital, stadium or factory.

Detection
Find, localize and track a behavior that is a trigger for a potential CBRNE incident evolving in a public transportation network (e.g., London metro underground system). Examples: person exhibiting behavioral signs consistent with certain diseases, or a person whose behavior (e.g., depositing a knapsack or suitcase on a train car or platform and then leaving the scene without it []).

Recognition
Refine the image(s) captured, using not only image processing techniques but the incorporation of semantic data and associative knowledge (e.g., GIS, CBRNE sensor data, text msgs, other non-specific camera data), and construct a template model for the target of interest's face and/or other body geometry.

Mining
Find match(es) for the target of interest (t-I) in databases of images and video (e.g., known or suspected terrorists, or persons exposed to H5N1 or XDR-TB). Search for prior locations of the t-I in the subway system, also scenes where there may have been persons in close and extended contact with t-I (e.g., meetings with conspiratorial associates, meetings with strangers who could have been exposed). Such findings would naturally trigger passing data to additional processing units that would be charged with following these tree-paths of relational and spatiotemporal interest.

Synthesis
Project forward to t-I next movements, route, destination, changes in appearance from different angles and viewpoints, different conditions of lighting, and different conditions of partial imaging as a result of being concealed in a dense crowd of people, a fast-moving train, etc.

Optimization
Refine what are the most likely of the possibilities with respect to images, routes, contacted people, prior itinerary and route before discovery, etc.

Some of the tasks that will high-demand in terms of computing (processing and memory) may include eigenvector functions, Morse and Ramsey models, PCA (principal component analysis), parallel decision tree traversing, parallel Bayesian evaluations, parallel neural network computations, METI (most evident topic of interest), and large numbers of conventional image processing segments especially those concerned with face recognition and object separation.

Health – I^3DIT™ []
Deformable image registration and in particular mutual information analysis [] is well-established within the field of medical imaging and in the control of radiation therapy (e.g., IMRT – Intensity Modulated Radiation Therapy) and now the emerging fields of nanoscale or molecular imaging and therapeutics. At the heart of the challenge requiring solid MPC is the need to perform large

inverse method computations such as are standard for wave scattering and in use for surface and subsurface imaging and acoustics. The particular challenge is to perform such tasks in real-time and for large data sets since the objective in early-stage detection of events at the tissue level or at the scale of molecular markers is to identify very small distinguishing features, such as a tumor that may be significantly less than 1cm diameter (approx.) or a concentration of antibody-tagged markers that have collected within a region of an organ []. The case model is drawn from the field of cancer imaging but can be applied to a very large number of similar medical tasks.

Detection
Find and localize in a non-stable sequence of image (sensor) frames collected from MRI, PET, or other methods a set of distinctive markers and perform comparisons with other image data originating from one or more of the following types of sources:

- prior images from the same patient using the same imaging modality
- alternative simultaneous or near-time imaging performed on the same patient
- historical pathological and normal imaging data from other patients

Recognition
Refine the image/sensor data and classify the targets as normal or pathological according to characteristic types. Fit the data set using attractor-diffusion and other methods (e.g., Morse-based volumetric separation) to realistic geometries.

Mining
Search through large sets of medical image (sensor) data streams to find similar or distinctively dissimilar objects and sequences of objects for use by experts in completing a diagnosis or a therapy plan.

Synthesis
Model how an identified and targeted data object could have looked under different conditions including any of the following parameter changes:

- earlier time, given certain assumptions on tissue vitality and growth
- projected time in the future, given certain similar assumptions
- hypothetical judgment about an interfering tissue surface
- adjustments for physical movement disturbances of the patient's body during the examination

Optimization
Project how a certain condition may be expected to evolve if treated with a certain pharmacological or radiative therapy.

Energy – I-Trans™ []
The problem of transportation in a world that is massively populated, with depleted energy resources, insufficient sustainable fuel alternatives, global climate changes exacerbated by carbon emissions, and overcrowding and inefficiency of streets and highways is a serious nightmare for our 21st Century. There are solutions that are economically sensible and that offer a balanced answer for the demands of private, convenient, individual-choice, on-demand transport (the automobile) and the needs to reduce sizes of vehicles, energy consumption and physical traffic loads. I-Trans™ is one such possible solution [], integrating the conventional automobile with limited-access urban "FlexCar" and "Clever" car and traffic concepts, plus innovative high-speed car-rail hybrid transport plans for intercity and long-distance travel, and intelligent traffic management that is based upon data acquisition and modeling that requires massive numbers of

sensors and inputs both onboard vehicles and along highways. This case model is very complex and what is suggested here can be divided into a number of MPC/HPC applications that serve smaller task components, each of which are reasonable topics for R&D and commercial deployment within the automotive and highway industries.

Detection
Find densities of traffic in particular areas and along routes. Find indications of congestion. Find areas of traffic activity such that there are indicators for upcoming traffic jams. More local to the individual car, find indicators of long-term travel being underway.

Recognition
Refine and complete by estimation patterns of sensor and onboard travel data from large numbers of vehicles and match them to simulation templates and historical patterns indicative of travel forecasts.

Mining
Search schedules and load estimates for alternative travel routes and alternative I-Trans methods, including: availability of local Clever-type commuter vehicles, availability of parking at various terminals, routes and schedules and available car-seats on vehicle-commuter trains (a key facet of the I-Trans concept), traffic estimates at select vehicle-commuter train debarkation points.

Synthesis
Planning routes and schedules and pre-load driving instructions for large numbers of individual drivers before and during their trips.

Optimization
Refine travel plans for individuals and trains on the basis of changing data such as accept/decline responses, new traffic and weather patterns, and other variables.

Lifestyle – "Orbis-Uno" – the Smart Healthy Safe Home
The home is increasingly complex because of the time demands upon people's lives including frequent and unexpected mobility, the proliferation of electronic technologies, and also changing lifestyle patterns including those for the disabled and elderly. The advent of new energy demands as well as opportunities to implement into the home new energy management protocols that require attentiveness and alertness in order for these systems to be both effective and safe is a further demand for introducing HPC solutions into the home in a manner that has not been often contemplated in the past as a realistic market for computing devices, even among internet-appliance and smart-home innovators.

Detection
Find and estimate if persons are in a given room and estimate the duration of future occupancy. Collect sensor data pertaining to motion (external and internal to the home), temperature, sound, occupant activities. Monitor E-Poly™ and other modular and non-centralized energy systems. Monitor water usage, air quality, and other parameters of interest. Coordinate the actions of multiple devices such as servo-controlled webcams in order to localize a person or activity of interest such as an elderly or infant member of the household.

Recognition
Identify the person in a given room and activity as a specific individual member of the household. Identify types of activities engaged by occupants of different rooms. Identify potential risk and

threat patterns. Create association-potential maps such as a linkage between a kitchen activity and an elderly or disabled adult or a young child, expected time intervals, and other parameters that could be used to provide notifications or warnings.

Mining
Search through known sequences and patterns of activity in order to identify anomalies that could merit special action and alerts. Search for rule-out models that will negate causes for concern and alert, because the problem of false positives is likely to be one of the largest hurdles to overcome with home-based tera-scale applications of this sort.

Synthesis
Extend association-potential maps to create expectations of household member activity that should be monitored at specific intervals or after a threshold of time, spatial movement, noise or silence, and so forth. Fill in the missing elements of an n-dimensional map of household activity by humans, animals and machines such that a larger pattern can emerge, such as one indicative of problems in locomotion by an occupant, presence of intruders such as robbers, appliances left on too long, water or wind damage during absence from the property, or non-optimal use of HVAC utilities.

Optimization
Estimate recommended courses of action for occupants, first responders, or energy management systems. One example from the energy domain is the estimation of balances between natural and artificial co-operative heating and cooling. Natural methods include opening of windows and use of skylights in stairwells. Artificial systems include solar panels and E-Poly™ sheets including draperies, circulating fans, hot water heating, and auto-controlled interior shades and doorway curtains.

Abstraction
From examples such as the above and others that have been developed experimentally by researchers at Intel and many other MPC research groups, it is possible to move toward an abstract model for DRMSO modeling that can serve in the long term to provide a toolset, not dissimilar to those developed for use-case system design and object-oriented applications building. The goal of such an abstraction process is simply to make life easier for everyone trying to use a multi-core architecture to solve problems that inherently have no clear bounds, no rigid structures, with respect to the types of input devices, pattern-recognition and data mining tasks, and ultimately the synthesis outputs that are expected to make a difference in an intelligent way for the end user.

What follows here are only some very elementary sketches for the steps that can be taken to build such an abstract machine. Any formalism must then be converted into some set of programming methods and ultimately functions that can be used, perhaps analogous to an API Library, by the application programmers.

Particularization

Adaptation and Assimilation

Notes and References

