Perspectives on the workshop series

Monique Werlen[a] and Denis Perret-Gallix[b]

[a] *LAPTH, BP 110, F-74941 Annecy-le-Vieux, France*
[b] *CNRS-JAPON, 3-9-25 Ebisu, Shibuya-ku, Tokyo 150, Japan*

Abstract. After a short history of the AIHENP workshop series started in 1990, we will sketch how, through the 7 meetings, we have been monitoring the computer revolution which is drastically transforming the way we perform basic research and, furthermore, how this series has been anticipating some of the major happenings. We will end up by proposing a critical review of the main goals of the workshop series and a short list of some current budding techniques and their possible applications to basic research.

SHORT HISTORY OF AIHENP

The ACAT'2000 workshop, 7th issue of the AIHENP series, marks the 10th anniversary of the series. The AIHENP acronym was meant to stand for Artificial Intelligence for High Energy and Nuclear Physics. But what does ACAT mean?: A CATapult to the future of Computing? A CATacombs for deadlocks? A CAThedral for the Holy GRID? . . . No, simply, "Advanced Computing and Analysis Techniques in Physics Research". This new acronym puts more emphasis on the profound and essential evolution of the data analysis techniques.

Six workshops have been organized until this one hosted by Fermilab. The first one was held in Lyon (1) (March 1990), followed by La Londe-les-Maures (2) (January 1992), Oberammergau (3) (October 1993), Pisa (4) (April 1995), Lausanne (5) (September 1996) and Heraklion (6) (April 1999). Each had its own personality and uniqueness thanks to the deepest involvement of the local organizers.

The goal of the first workshop in the city hosting the French High Energy and Nuclear Physics computing center (CCPN) was to focus on technologies born from the conceptual methods developed in the field of Artificial Intelligence and to discuss their uses in high energy and nuclear physics research. "AI" was selected in a slightly provocative way to attract the attention on our willingness to confront real applications to the sixties smoking dreams. But the needs for more embedded autonomy and adaptability in the software were already a reality. With the beginning of the big LEP experiments, the implementation on on-line real-time and on off-line data analysis software of more intelligence was becoming quite inescapable.

Essentially three sensitive fields had been identified:

- Software engineering: In French: "Génie Logiciel" but "génie" also means genius which should go well along with artificial intelligence! More seriously, at this time, only simple text editors were available. They were providing no support for the large software developments needed for the on-line and the data analysis of those long term and complex experiments. For the first time, in our field, it was made clear that some sort of "intelligent" support was necessary for the software development of these multi platform, multi language, multi user, multi developer, multi country undertakings.

- AI algorithms: A few people were experimenting how basic research could benefit from the implementation of the most promising AI developments. Here the field was quite mined, due to the poor delivery of AI, but we will see how the focus of interest moved from the nineties to the new century and how, nowadays, some of those algorithms are central to fore front physics research.

- Symbolic algebraic Manipulation and automatic calculations: The computation of signal and background processes for LEP and the extensive need for multi-purpose event generators had started the development of automatic calculation based on algebraic languages or on diagrammatic numerical computation. This was the first meeting, world wide, of the different groups involved in this now burgeoning technique.

It was also quite clear, 10 years ago, that the more the computing technology was progressing, the more the plain physicists were at pain in this intrusive jungle with two serious problems that we haven't fully solved yet:

- A loss of manpower in the software development: Many physicists could not efficiently contribute to the software development as they did not have the basic knowledge of the new programming paradigm.

- A loss of physics vista on the computer programs: The computing experts were losing ground on the physics matters for being too much involved on the programming technicalities.

So this series was seen as get-together of experimentalists, theorists and computer scientists for hand to hand presentations and discussions expecting to soften the pain of some of us and bring perspectives to others.

Monitoring the computing revolution

Back in the seventies, high energy experiments were being performed without computers. The data taking was essentially performed by reading a set of scalers, counting particles crossing cleverly arranged detectors gated by triggers based on the timing coincidence of fast devices. The data analysis was done through the interpretation of histograms drawn by hand representing the scaler contents. Final results like cross-section were calculated by applying geometrical detector acceptance and dead-time correction factors.

No computers ... life was easier ... more time to think ... physicists were doing physics. Why did we have to introduce computers in this wonderland story? Because they were there ! or at least the first mini-computers were being assembled. The industry was looking for guinea-pigs to launch the largest ever revolution since the wheel (see the workshop series Logo). Is this actually true ? No, we do not think so, we think that Mother Nature was even more provoking in this happening.

Nature was requiring a push to higher energy interactions; although looked for, the intermediate boson, later on called Z, was not found at 29 GeV on a fixed target experiment! Higher energy collision studies could only be achieved by colliders and 4 π detectors which could not anymore be registered by scalers. Furthermore, the higher the energy the more complex the final state, the triggering and detection system. This did drive a complete change in the event read-out and recording technique. Thanks to the newly available micro-electronics and computers technologies, the modern HEP could actually set in.

But, in passing, that the way, in a more general stand, basic research progresses: It attempts to approach the Truth by glancing in the unknown on top of the shoulders of the latest technological achievements.

Since 1990, when the first workshop of the series was organized, we have seen an evolution in the main topics of the workshop, a sliding of the main focus of the sessions.

In the software engineering, at first, we had descriptions of CASE systems and the technicalities of the software design, presentations of "new" languages (Ada, Prolog, Eiffel) and an introduction to interactive analysis.

Then, the focus moved on to the object-oriented paradigm and the starting of the web, then C++ became the center of most of the developments, finally parallel computing got a lot of attention due to the exploding CPU power needs and this year the GRID, the world-wide net-computation entered in the dance.

In AI, the emphasis was first put on the expert systems for the diagnosis of the hardware equipment, but rapidly the main buzz words became neural-networks and genetic algorithms. This session due to the overwhelming interest in neural-nets turned out to be the largest session of the workshop in terms of contributions and attendance. This year we were shown that the latest algorithms are now so powerful that even Higgs can be "detected" by neural-nets (7)!:)

In symbolic manipulation, we started with descriptions of symbolic manipulation languages and their ability to mimic mathematician activities. In addition, some pioneering work on automatic calculations were presented. Then, advanced calculations (1-loop) and physics results produced by automatic procedures were shown. Finally, the question of dealing with those lengthy computations raised the need for more advanced computer technologies like automatic formula optimization and parallelism. Other uses of symbolic languages were exemplified for computing super-algebra string oriented theories.

Anticipating the wave

It is always comforting for the organizers of a workshop to "discover", even if it went confidentially, that some major event did actually occur. For example in January 1992, in La Londe-les Maures, Tim Berners-Lee, associated with J.F. Groff, R. Cailliau and B. Pollermann, presented the " World-Wide Web: An information Infrastructure for High Energy Physics" (8) probably one of the very first presentations of the WEB in front of a wide audience (250 people).

Sometimes new ideas or technologies have a hard time imposing themselves. This has been the case for the automatic calculation of HEP processes. Although most of the symbolic manipulation language designers were high-energy physicists, it was not recognized by the rest of the community that those techniques although not "replacing" the theorists, could alleviate some of their calculation burden. At best this was considered as a learning tool for beginners. It has been therefore a big pleasure to acknowledge the participation in this series of those famous designers. G. Gonnet, A.C. Hearn, J. Vermasseren, S. Wolfram and even M. Veltman (9) (1999 Nobel prize winner) have in some way been backing up this activity. These workshops have contributed to the takeoff of the automatic computation systems which are now fully embedded in the event generators used at colliders.

As already mentioned, during these last 10 years, one has seen the rising importance of neural-net selection in data analysis and in on-line triggering. This workshop has served as a permanent forum for the development of these new techniques.

BACK TO THE FUTURE

What will be the future of this series? Well, what you, participants, will make of it. The Advisory committee will be keen at enforcing the basic goals which are the foundation of this series, but the content, the "substantifique moelle", you, the actors on the field, you will make it happen. We would like to hear from you any proposals or ideas to help us improve this series.

Panel sessions for in depth discussions

This is an efficient way to provoke discussions between people. We all did appreciate the panel sessions organized at ACAT'2000, regretting maybe that they were in parallel sessions. We intend to have more of those in the next workshop.

A catalysis of ideas and concepts

A single idea is rarely good by it-self, it finds its real application through its synergy with other concepts. Let's take the WEB story as an example. Everything started with the idea of creating hyper-link between documents on different computers, but the full power of the WEB came with the development of the graphic interface. We remember using the very first text version of the WEB at CERN, we were not convinced that this tricky way of typing numbers to refer to such and such keywords would ever initiate the so-called Internet Revolution.

What are the new ideas now popping up and about which we would like to have discussions in the future issues of this series? Well this is not easy to say and others in this workshop will address this question, but if we can add our pinch of salt, here is what we personally would like to track in the current developments.

- Virtual Reality: So much have been said about it in the simulation and game playground. But what can this new technology bring to us, actually? First of all this is really the next step in event scanning. The ability to manipulate a complex HEP event as we would with an object of "art", "going" to the interaction point and by Lilliputian steps inspecting all secondary vertices or looking in the particle trajectories using our personal, human, fuzzy logic to estimate, "by eyes", the momentum of the particle or to identify and correct wrong pattern matching, is probably a big plus. This technology can also be useful for the maintenance of the huge electronics and computer environment of the LHC experiments. Expert Systems have been of a great help already to send alarms and to indicate the series of actions to be performed. But imagine wearing those goggles and gauntlet and being guided directly by the computer, like a jet pilot to its target, to the right red button on the module 16 of the bay 15 in the rack 2587 of the room 12 to do the correct manual reset needed by the on-line to resume ... a plus no ? More applications include the help in the design of the detector by visual inspection, the simulation and the guidance for on site detector repairs .... But additional developments are probably needed to realize efficient and fully operational systems.

- Nanotechnology, molecular machines, nanobots are fashionable and strong expectations are put on this budding technology. Matter will become software, and vice versa (we would say). "So some day soon, we could download hardware from the Net just like we download software today." (10) The ability to build molecule by molecule (or atom by atom) electronic devices just as one writes on an magnetic disk by arranging magnetic grains would make possible the building of specific circuit dedicated to a very precise application. It would be possible to create "in hardware" the software algorithm most appropriate for a triggering signal and take the full benefit of the orders of magnitude faster execution time. Moreover this could be built on-line, on demand, when needed. The gap between software and hardware would vanish, as all software could turn into hardware in the running of the program. This give just a taste of what could be the next revolution, but needless to say that many workshops will go before those ideas will be transformed into down to earth applications.

Table 1. Comparison of performances for the salesman problem with 75 cities

algorithm	result	nb of iterations
Particle Swarm algo.	535	3480
Genetic algorithm	545	80000
Evolutionary programming	542	325000
Simulated Annealing	580	173250
Optimal solution	535	

Source: Dorigo and Gambardella, BioSystems 43,73 (12)

- Swarm intelligence, bio-genetic research, ...: Those optimization algorithms are based on the observation of the social behavior of insects (11). An ant colony selects the shortest path to a target thanks to the deposited pheromone. It is actually an example of collective problem solving as opposed to centralized control. Salesman problem algorithms based on this principle give quite good results (12) (see table 1). Traffic routing for Internet packet or transportation vehicles issues can also be addressed by this algorithm. But most interestingly, the particle swarm algorithm (13) can be used in place of the learning algorithms needed by the neural-nets and may evolve both the weights and the structure it-self of the net. Last, but not least those systems are well suited for parallel implementation.
- New Internet Protocol: New Internet protocols like IP v6 are being discussed, increasing dramatically the number of addresses. This would make possible for each important part of an experimental detector to have a different IP address and therefore to embed its own local intelligence wherever it would be located. This is an important aspect when building widely distributed intelligence for an electro-mechanical equipment.

Promoting new initiatives

The workshop is also a place for informing the participants of new important programs. The GRID project has been largely discussed this year, but other projects will be discussed in the future including Internet II, a high speed Internet dedicated to R&D labs and universities.

Establishing Standards

We would like to advocate here, again, the need to establish standards. Although standards are not always possible or even desirable for some developments, we have to do our best to reach minimum agreement on common standards every time it is possible, just not to re-invent the wheel (especially if you happen to make it squared). Look at the mess of maintaining the same C/C++ package on different computers with different compilers. The big successes in the software history are the cases where many people have contributed to develop, to enhance and to extend the package. This cooperative development has been made possible thanks to a clean definition of the basic standards. Standards may come 'de facto' after long test and trial or may be the output of committees and again they take time to be set. But one has to go through these painful steps to benefit later from the well designed basics. We have to learn to think globally, collectively and cooperatively from the start of the package design.

Open to all fields of research

Although this workshop series started with the application field built in in its acronym (HENP: High Energy and Nuclear Physics), we soon opened it to other fields of research, first in physics, astrophysics and then even more widely to space science, telecommunication, biology. This has never been really successful in the sense that the majority of the speakers and of the attendance always did come from HENP. It is therefore a pleasure to acknowledge this year the participation of astrophysicists and particle accelerator physicists. But in future issue, we would like to have it even more attractive to other research fields as long as they are facing similar problems or using similar approaches or because they bring to particle physics new view points. For example:

- the need for large computing power requiring the use of supercomputers or large scale farm-like systems is found on many problems: in molecular modeling, bio-computing, 3D animation or Internet queries.
- the handling of huge amounts of data like for the LHC experiments is also critical for the human genome or the Internet storage.
- the security and reliability of the computer systems are key points in all systems.

Different research fields like meteorology, fluid dynamics or bio-computing may use similar algorithms for the simulation of processes, the parameter optimization or the multi-dimensional integration. Moreover, we in HEP can learn from discoveries made in other domains like in the brain research or, as we have already seen, from the study of the social behavior of ant colonies.

Permanent link

Finally, we want this workshop to be a permanent link between developers, users and the industry through:

- email, newsgroups and a decent WEB page (this will be done in the coming month $http://www.lapp.in2p3.fr/aihep/$)
- Intermediate video or net conferences
- Tutorials
- Products show-room and demos

ACKNOWLEDGMENTS

The ACAT'2000 has been without contest a really enjoyable and exciting meeting where many new ideas have been discussed and, we hope, new collaborations have been formed. On behalf of the International Advisory Committee, we want to express our deepest thank to Pushpalatha Bhat, Matthias Kasemann and to all the members of the local committee for this smooth organization.

Several proposals have been put forward for the organization of the next workshop. We are happy to announce that the international advisory committee has enthusiastically accepted the Moscow application. We wish our Russian colleagues the same success in 2002.

REFERENCES

1. *New Computing Techniques in Physics Research. Proc. 1st International Workshop on Software Engineering, Artificial Intelligence and Expert Systems in High Energy and Nuclear Physics Lyon (March 19-24 1990)*, ed. by D. Perret-Gallix and W. Wojcik, CNRS, Paris 1990.
2. *New Computing Techniques in Physics Research II. Proc. 2nd Int. Workshop on Software Engineering, Artificial Intelligence and Expert Systems in High Energy and Nuclear Physics, La Londe-les-Maures (Jan. 13-18 1992)*, ed. by D. Perret-Gallix, World Scientific, Singapore 1992.
3. *New Computing Techniques in Physics Research III. Proc. 3rd Int. Workshop on Software Engineering, Artificial Intelligence and Expert Systems in High Energy and Nuclear Physics, Oberammergau, Bayern, Germany (Oct. 4-8 1993)*, ed. by K.H. Becks and D. Perret-Gallix, World Scientific, Singapore 1993.
4. *New Computing Techniques in Physics Research IV. Proc. 4th Int. Workshop on Software Engineering, Artificial Intelligence and Expert Systems in High Energy and Nuclear Physics, Pisa, Italy (Apr. 3-8 1995)*, ed. by B. Denby and D. Perret-Gallix, World Scientific, Singapore 1995.
5. *New Computing Techniques in Physics Research V. Proc. 5th International Workshop - AIHENP'96 Lausanne (Sep. 2-6 1996)*, *Nucl. Inst. Meth.* **A389 (1-2)** (1997), ed. by M. Werlen and D. Perret-Gallix.
6. *New Computing Techniques in Physics Research VI. 6th Int. Workshop on Software Engineering, Artificial Intelligence and Expert Systems*, ed. by G. Athanasiu and D. Perret-Gallix, Parisianou S.A., Greece (2000).
7. ALEPH Collaboration, "Observation of an excess in the search for the standard model Higgs boson at ALEPH", CERN-EP/2000-138, *Phys. Lett.* **B** in press. Bhat P.C., Gilmartin R. and Prosper H.B., "Strategy for discovering Higgs at the Tevatron", *Phys. Rev.* **D62**, 074022 (2000).
8. Groff J.-F., Berners-Lee T.J., Cailliau R., Pollermann B., "World-Wide-Web: An Information Infrastructure for High-Energy Physics", in (2), pp. 157-164.
9. Veltman M., "Symbol Manipulation: Early History and Present Prospects", in (3), pp. 451-458.
10. Ellenbogen, J.C., "Matter as Software", presented at the *Software Engineering and Economics Conference, McLean, 2-3 April 1997*; Gerd Binnig interviewed by Otis Port in Business week on-line 99_35.
11. Bonabeau E., Dorigo M. and Theraulaz G., *Nature* **406**, 39 (2000).
12. Dorigo M. and Gambardella L.M., *BioSystems* **43**, 73 (1997).
13. Kennedy J. and Eberhart R., "Particle Swarm Optimization", in *Proc. of the 1995 IEEE Int. Conf. on Neural Networks, Perth* IV pp. 1942-1948.