DD/

CERN - ECP 91-20

EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

See 9151

CERN/ECP 91-20
24 September 1991

9

CM-P00056570

**AN ENTITY RELATIONSHIP APPROACH TO THE EVENT
AND DETECTOR MONTE CARLO SIMULATION
IN HIGH-ENERGY PHYSICS**

LAA Project

L. Cifarelli^{1),2),3)}, G. La Commare³⁾ and M. Marino^{3),4)}

CERN, Geneva, Switzerland

- 1) University of Naples, Italy
- 2) INFN, Bologna, Italy
- 3) European Organization for Nuclear Research, Geneva, Switzerland
- 4) World Laboratory, HED Project, Lausanne, Switzerland

Abstract

A new system has been implemented for Monte Carlo simulation purposes in the framework of CERN/LAA-MSL Project, to deal with the ever-growing complexity of event and detector simulation in high-energy physics. The system, called FMC (Full Monte-Carlo Chain), comprehends three independent subsystems to define particles and decay properties, PDKDB (Particle and Decay Data Base), to describe the detector configuration, Set-Up Descriptor (SUD), and to deal with the event kinematics via interfacing routines among various Event Generators (EGs), the MEGA (Monte Carlo Event Generator Adaptor). An extensive use of the Entity Relationship (ER) techniques has been made to model the data with particular care on consistency and compatibility. By integrating the system to the GEANT tracking program, a complete simulation chain is obtained.

Communication to the Società Italiana di Fisica, LXXVII Congresso Nazionale,
L'Aquila, Italy, 30 September-5 October 1991

1. INTRODUCTION

The scale and complexity of High-Energy Physics (HEP) experiments is continuously increasing. Physicists are confronted with simulation studies which require more and more care; the amount of data to store, interpret, manage, and control calls for specialized systems to easily and efficiently accomplish these tasks.

The present situation can be roughly depicted in the general Monte Carlo simulation framework as organized into the following steps:

- description of the experimental set-up in terms of geometrical information, relevant properties of materials used within detectors, and detector responses;
- management of particle properties and decay modes for event generation and tracking inside the detector, taking into account all physical effects due to the particles themselves, their interaction with the matter, and with the magnetic field if any;
- interpretation of event kinematics to be processed during the simulation.

Usually the information is stored, into internal data structures of the simulation program(s) used, at initialization time: the user is therefore required to fulfil a preliminary, often long and tedious, input phase.

In this work we present a new system, the so-called Full Monte Carlo Chain (FMC) [1], to manage this information; moreover, we provide an automatic input interface between the most widely-used QCD Event Generators and the well-known GEANT tracking program [2]. Some of the ideas implemented herein have been developed in collaboration with the ZEUS Group of Bologna University, in particular for the part concerning the description of the detector.

2. THE ER APPROACH TO DATA MODELLING

Most of the complication in handling a great amount of data comes from a poor use of data modelling. Information is often inserted into *ad hoc* common blocks with different organizations in different parts of the programs. As a direct consequence, the relationships among data are difficult to handle, and this leads to a coding which easily expands in a rather confusing way.

A great help comes from Data Base Management Systems (DBMSs) capable of easily handling huge amounts of data, still preserving all their essential mutual relationships. The DBMSs generally assume a specific model to describe data. Usually, more time is to be devoted to the design phase, but a careful organization of the data structures is of great advantage: programs can be written in a much clearer way, which translates into a simpler coding. The natural consequences of this approach are an easier debugging phase, and a better control on future program developments.

Following these lines, we have implemented in the framework of the CERN/LAA-MSL Project [3] a system to handle the different steps needed to develop a full Monte Carlo simulation program for HEP experiments; it is composed by three subsystems to deal with:

- the Event Generators (EGs) which provide the physics input,
- the geometrical description of the experimental set-up, and the definition of the response from its sensitive detectors,
- the management and control of particle properties and decay modes.

To develop our solution, we made extensive use of a FORTRAN implementation of the Entity Relationship (ER) model, the ADAMO system [4], embedded with a suitable set of routines to manipulate the data, make consistency checks, retrieve information, etc. The ER model represents the world by means of entity sets, their attributes, and their mutual relationships.

The advantage of proceeding along this path relies on the fact that once we achieve a complete data description, the complexity of the coding to interact with the information is highly reduced, as well as the effort to generate it.

3. THE SYSTEM

The proposed FMC system, is constituted by three independent subsystems dedicated to different kinds of data: physics event MEGA [5] (Monte Carlo Event Generator Adaptor), particle and decay properties (PDKDB) [6], tracking parameters and detector description SUD [7, 8] (Set-Up Descriptor), as illustrated in fig. 1.

Fig. 1 FMC data flow diagram

The MEGA package is interfaced with the main EGs (LUND [9], HERWIG [10], EUROJET [11, 12], ISAJET [13]) in a transparent way: the user should only indicate where data are stored, and which EG produced them.

From inside FMC the user can modify data relative to two kinds of particles: the heavy flavoured (HF), and long-living (LL) particles.

All the particles are usually associated to the beam interaction vertex. If a detailed description of secondary vertices is needed, FMC generates decay vertices from HF decay. The HF decay products are then linked to the proper secondary vertex before being transferred to GEANT for tracking.

Long-living particles, because of their long flight, can be influenced by the presence of a magnetic field, if any, and by the interaction with the matter. Their behaviour is, in general, related to the experimental set-up. The FMC lets these particles decay inside GEANT so as to generate their decay products under the detector influence.

The detector description and the particle information normally require a substantial action from the user: data need to be revised, improved and edited frequently before a final configuration is obtained. An efficient and user-friendly approach has been aimed at by developing two interactive subsystems (SUD and PDKDB), based on a command interpreter, where the user can access the data in a transparent way. A self-explanatory dialogue method, together with an on-line help facility available for each command, enables the user to easily define the data and interact with them. Commands are grouped in Object/Action menus. At any moment the session can be saved by updating an existing data file, or by creating a new one.

Among the many possible ways of describing a detector, the one already developed and tested in the GEANT program has been adopted. The SUD can be divided in two basic parts:

- the description of the detectors,
- the definition of detector responses,

which allow the set-up specification.

The basic element to build a detector is a volume (and its possible divisions), which is characterized by its shape and dimensions. The volume is positioned and oriented in space with respect to a given reference frame. A tracking medium (made of one or more materials) is assigned to the volume. Among the volumes, the sensitive ones give a signal when traversed by a particle. The response of the volume is described in terms of hits and their possible digitizations. In summary, we have the following list of actions:

- define a set of volumes and their divisions (if any),
- define their positions and orientations,
- assign a tracking medium to each volume,
- define the sensitive volumes and their hits/digits,
- assign the generated detectors to one or more set-ups.

This way of describing an experimental set-up fits well in an ER model. The last subsystem, PDKDB, handles data on particles and particle decay channels; it contains a set of initial data that are accessed automatically at initialization time. These data are obtained from the Review of Particle Properties, and from well-known EGs (e.g. EUROJET, LUND, etc.). Data are updated by means of a dedicated data file, regularly released by the Particle Data Group [14]. The subsystem ensures the data compatibility and consistency through an exhaustive set of checks. Data are managed by two separate structures: the Kernel and the Delta data base. Kernel holds the initial information on

particles and decay modes. Delta handles new data and modifications. The separation of information leads to a full robustness of the entire data base. In fact, Kernel data are not directly accessible to any operation. However, from the user standpoint, the interaction is operated directly with the Kernel.

By internal convention, only the particle of each particle/antiparticle pair is present in the data base. Properties and decay modes of antiparticles are automatically obtained through an internal charge conjugation.

Operations are divided in three groups, relative to particles, decay channels, and I/O. Particles can be inserted, modified, removed, selected for GEANT simulation; their parameters can be retrieved according to various selection options chosen by the user. Operations are checked to be consistent with the decay channel information relative to each particle.

Operations of insertion, modification, removal, list, forcing, and temporary disabling are foreseen for all decay channels. A check operation on total branching ratios is required by the system to fix the data base contents; an automatic scaling, to keep the mutual proportionality between channels, is available.

Input/output operations allow the user to physically store the information contained in the data base. The entire data base (Kernel plus Delta) or the session modifications (Delta) can be saved either in ADAMO format, or in a sequential file, either formatted or unformatted.

4. THE INTEGRATION WITH GEANT

The subsystems presented herein are application-independent, they can be used by any HEP program either in the input phase, or as data bases embedded in the coding, using the ADAMO packages [4].

The system is available in a VAX-VMS environment, nevertheless each subsystem produces machine-independent files that can be used in any platform where ADAMO is installed.

The interface to GEANT is obtained through calls to FMC routines inserted in the GEANT user routines. This will produce an automatic transfer of the information to the proper GEANT data banks. The integration foresees a modular interface with the system: the user can use one or more of the three data bases in his own GEANT application by setting the proper value to the system data cards, one for each subsystem, leaving unmodified the rest of the application.

5. CONCLUSIONS

A new system to describe data for Monte Carlo HEP programs has been developed, leading to a simple management of the information. A particular application of the system has been implemented in connection with widely-used QCD Event Generators and the GEANT tracking program, showing that it is indeed possible, for future HEP experiments, to go towards a new way of handling huge and complex data sets.

REFERENCES

- [1] C. La Commare and M. Marino, FMC: Full Monte Carlo Chain user guide, Technical Report CERN/LAA-MSL 91-06, version 2.0 (April 1991).
- [2] R. Brun et al., GEANT3 user's guide, CERN/DD-EE 84-1 (September 1987).
- [3] L. Cifarelli, Super Monte Carlo Simulation at 16, 40, 200 TeV, Conf. on Physics up to 200 TeV, Erice, Italy (July 1990), ed. Plenum Pub. Co., London, proc. of the International School of Subnuclear Physics, 28th Course (to appear).
- [4] S. Fisher and P. Palazzi, The ADAMO data system programmers manual, version 3.1, ADAMO Note (June 1991).
- [5] F. Anselmo et al., MEGA: Monte Carlo Event Generator Adaptor, Particle World, vol. 2/4 (1991).
- [6] G. La Commare, M. Marino and B. van Eijk, PDKDB: particle and decay channels data base, CERN/ECP 91-19 (CERN/LAA-MSL 91-8) (September 1991) to be submitted to Particle World.
- [7] A. Bassi et al., Detector simulation for complex multi-TeV experiments, proc. Conf. on New Technologies for Supercolliders, Erice, Italy, September 1990, ed. Plenum Pub. Co., London, and CERN/ECP 91-18 (1991).
- [8] G. La Commare and M. Marino, SUD/IS (Set-Up Descriptor/Interactive System) User Manual, version 1.0, CERN/LAA-MSL 91-12 (June 1991).
- [9] The LUND Monte Carlo programs, CERN/DD pool programs W5035/W5045/W5046/W5047/W5048 long write-up (April 1987).
- [10] HERWIG, CERN/DD pool program W5037 (October 1990).
- [11] B. van Eijk, EUROJET: a Monte Carlo program for JET simulation, Technical Report CERN/EP 85-121, presented at the 5th Topical Workshop on Proton-Antiproton Collider Physics, St. Vincent, Italy (1985).
- [12] A. Ali, B. van Eijk and I. ten Have, Heavy flavour production in large transverse momentum proton-antiproton collision, Nucl. Phys. B292(1) (1987).
- [13] ISAJET, CERN/DD pool program W5036 (October 1990).
- [14] F. Carminati et al., Information modelling for Monte Carlo Event Generators, eds P. Palazzi, R. Brun and P. Kunz, proc. of the 14th INFN Eloisatron Project Workshop on Data Structures for Particle Physics Experiments — Evolution or Revolution? Erice, Italy (November 1990), World Scientific Publ. Co., Singapore.