# An Ontology Design Pattern for Particle Physics Analysis.

David Carral,[1] Michelle Cheatham,[1] Sünje Dallmeier-Tiessen,[2,3] Patricia Herterich,[2,4] Michael D. Hildreth,[5] Pascal Hitzler,[1] Adila Krisnadhi,[1,8] Kati Lassila-Perini,[6] Elizabeth Sexton-Kennedy,[2] Charles Vardeman,[5] Gordon Watts[7]

Wrigth State University,[1] CERN,[2] Harvard University,[3] Humboldt-Universität zu Berlin,[4] University of Notre Dame,[5], Helsinki Institute of Physics[6], University of Washington,[7] University of Indonesia[8]

**Abstract.** The detector final state is the core element of particle physics analysis as it defines the physical characteristics that form the basis of the measurement presented in a published paper. Although they are a crucial part of the research process, detector final states are not yet formally described, published in papers or searchable in a convenient way. This paper aims at providing an ontology pattern for the detector final state that can be used as a building block for an ontology covering the whole particle physics analysis life cycle.

## 1 Introduction

Particle Physics, the study of the fundamental building blocks and forces of our universe, involves some of the largest experimental apparatus ever constructed, like the ALICE, ATLAS, CMS, and LHCb experiments located at the Large Hadron Collider (LHC) at CERN. Each of these "experiments" is a very large collaboration of physicists who work as a team to design, build, and operate the particle detectors and to produce measurements characterizing the particles that make up the universe. The measurements are inherently statistical in nature: often billions or trillions of particle collisions are analyzed to determine probabilities or probability densities associated with a given physical process. Because many of the experiments collect multi-purpose data, careful attention must be paid to defining the measurement that is to be made.

Despite the many thousands of papers published since the advent of particle physics in the 1940s, the field has no formal way of representing or classifying experimental results – no *metadata* accompanies an article to formally describe the physics result therein. A number of scenarios would be enabled with such a representation. For example, a physicist from ATLAS, or a theorist, could search an external database for previous work done by CMS in order to compare results. Even a physicist inside ATLAS could search an internal database for previous examples similar to a planned analysis; a substantial amount of time and effort can be saved by starting from some preexisting work.

We intend to address this situation with our ontology design pattern. Results in particle physics take many forms, but all are based on the selection of a target set of characteristics, a *detector final state* that defines the ingredients of the measurement. The fundamental unit of particle physics is the individual interaction of a set of particles, or an "event." An event could, for example, be captured from a single interaction of counter-rotating particles in a collider or from the collision of a high-energy cosmic ray in the atmosphere. The selection characteristics refer to properties of an event and can describe the presence or absence of specific particles observed by the detector in the aftermath of the collision, or potentially more global properties of the products produced in the collision, such as the total energy released. Since the physics results we wish to describe and preserve in a repository are all based on the selection of *one or more* detector final states, this is a necessary ingredient of an ontology covering the whole particle physics analysis life cycle.

Competency questions have been recognized as a good approach to detect and generalize the modeling requirements from multiple domains that an ontology can represent. They are queries that a domain expert would be expected to run against a knowledge base. For the proposed final state ODP, such competency questions include:

1. Retrieve all analyses requireing particles to have an invariant mass near the Z pole.
2. Retrieve all analyses that used jets in the final state.
3. Retrieve all analyses that veto extra leptons.
4. Retrieve all analyses requiring large missing energy.

## 2 Formalization

This section presents the detector final state pattern by discussing the more interesting classes, properties, and axioms. Description Logics (DL) [2] notation is used to present the axioms. To encode the pattern, we make use of the logic fragment $\mathcal{SROIQ}$ as defined in [4], which is the basis for the OWL 2 DL standard [3]. The proposed ODP has been formally encoded using the Web Ontology Language (OWL).[1] A schematic view of the pattern is shown in Figure 1.

**DetectorFinalState:** A detector final state (DetectorFinalState) formally describes and structures information about a physics analysis (measurement) that is defined by its use of a common set of particle physics characteristics. As such, it must describe those characteristics of the fundamental "event" that have been selected to make the measurement. It is defined, amongst other features, by a set of particles/objects (PhysicsObjects) contained in the event and by global quantities formed by performing some operation on the ensemble of objects contained

---

[1] The pattern can be downloaded from `www.dropbox.com/sh/0upr45j1awd4q0d/AAAw9BQ2eZIWBIh_rBpP1Uu1a?dl=0`.

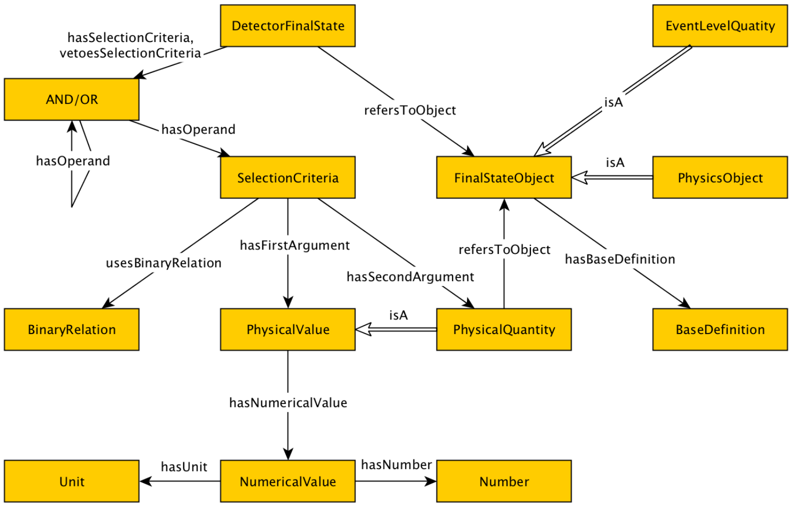

**Fig. 1.** A schematic view of the DetectorFinalState ODP

in an event (EventLevelQuantity). Both particles/objects and the ensemble measurements are referred to in this pattern as final state objects (FinalStateObject).

With the following axioms, we enforce that (1) every detector final state must refer to at least final state object, (2) all final state objects are either event level quantities or physics objects and (3) all event level quantities and physics objects are final state objects.

$$\mathsf{DetectorFinalState} \sqsubseteq \exists \mathsf{referstTo}.\mathsf{FinalStateObject} \tag{1}$$

$$\mathsf{FinalStateObject} \sqsubseteq \mathsf{EventLevelQuantity} \sqcup \mathsf{PhysicsObject} \tag{2}$$

In order to select events of interest, these objects are subject to selection criteria that are used to define a collection of events that serves as the basis for a physics measurement, and hence must be captured in the pattern. A detector final state (DetectorFinalState) then conveys numerical information describing the selection. This numerical information is referred as the selection criteria (SelectionCriteria) which models a complex boolean set of unary and binary restrictions. We make use of the classes And and Or to define complex selection criteria.

$$\mathsf{DetectorFinalState} \sqsubseteq \exists \mathsf{hasSelectionCriteria}.(\mathsf{SelectionCriteria} \sqcup \mathsf{And} \sqcup \mathsf{Or}) \tag{3}$$

$$\mathsf{And} \sqsubseteq \exists \mathsf{hasOperand}.(\mathsf{SelectionCriteria} \sqcup \mathsf{And} \sqcup \mathsf{Or}) \tag{4}$$

$$\mathsf{Or} \sqsubseteq \exists \mathsf{hasOperand}.(\mathsf{SelectionCriteria} \sqcup \mathsf{And} \sqcup \mathsf{Or}) \tag{5}$$

**FinalStateObject:** As mentioned above, there are two different types of final state objects in our model: physics objects and event level quantities. Each of these is defined by a restricted vocabulary and will point to another class, namely BaseDefinition, which will serve as a hook to provide more specific information about these types. Axiomatically, then, every PhysicsObject and every EventLevelQuantity are FinalStateObjects:

$$\mathsf{PhysicsObject} \sqsubseteq \mathsf{FinalStateObject} \tag{6}$$

$$\mathsf{EventLevelQuantity} \sqsubseteq \mathsf{FinalStateObject} \tag{7}$$

In order for these quantities to have meaning, each of the FinalStateObjects *requires* a BaseDefinition that describes the criteria for the creation of the FinalStateObject.

An example typical selection could be "retrieve all detector final states involving some electron with $p_T > 40$ GeV." In this case, the selection requires a particular type of final state object and has a restriction of 40 GeV, a PhysicalValue with a NumericalValue of 40 and a Unit of GeV, on the PhysicalQuantity $p_T$, which is shorthand for the momentum of a particle in the plane transverse to the beam axis. In general, a SelectionCriteria must indeed have a FirstArgument specifying a value and a SecondArgument specifying of which PhysicalQuantity this value is an instance, with some sort of binary operator ($<$ or $>$, for example) specifying the desired relationship.

For more information as to how information is stored using the pattern see `www.dropbox.com/sh/0upr45j1awd4q0d/AAAw9BQ2eZIWBIh_rBpP1Uu1a?dl=0`. We not only include terminological axioms in our ontology but also populate the pattern using data from existing publications [1].

## 3 Conclusions and Future Work

This paper proposes a generic ODP to capture the common core of experimental results from particle physics research. More specifically, it provides a precise description of a detector final state which can be used to assign meaningful metadata to the output produced by LHC. In future iterations we plan to extend axiomatization and populate it using real-world data to validate its usability.

## References

1. Aad, G., et al.: Search for pair-produced long-lived neutral particles decaying in the ATLAS hadronic calorimeter in *pp* collisions at $\sqrt{s} = 8$ TeV. Phys.Lett. B743, 15–34 (2015)
2. Baader, F., Calvanese, D., McGuinness, D., Nardi, D., Patel-Schneider, P. (eds.): The Description Logic Handbook: Theory, Implementation, and Applications. Cambridge University Press, second edn. (2007)
3. Hitzler, P., Krötzsch, M., Parsia, B., Patel-Schneider, P.F., Rudolph, S. (eds.): OWL 2 Web Ontology Language: Primer. W3C Recommendation (27 October 2009), available at `http://www.w3.org/TR/owl2-primer/`
4. Horrocks, I., Kutz, O., Sattler, U.: The even more irresistible $\mathcal{SROIQ}$. In: Proc. of the 10th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR 2006). pp. 57–67. AAAI Press (2006)