Extract of an e-mail from Daniel Froidveaux

I) Current status (to the best of my understanding), see
https://twiki.cern.ch/twiki/bin/view/Atlas/InDetTruthAlgs for more information)

- RDOs for ID are produced after digitisation step with associated RDO truth objects (this is not documented in the above link and I am not sure where to find this aside from browsing the code)

- The RDO truth objects contain all the necessary information (the list of all tracks having contributed to the particular RDO with the energy deposited for each of these tracks)

- For the purpose of documenting things as accurately as possible for this review and for work, which will follow on from it, it seems to me important to write down for each subsystem what the RDO and RDO truth structures are (they should perhaps sometimes be different, as illustrated below e.g. for the SCT and MDTs):

1) The SCT RDOs are neither individual strips nor real clusters, they mirror faithfully the current SCT ROD readout. Example: in a given module, consecutive strips n to n+4 are hit. There will be three RDOs, one with strips n and n+1, one with strips n+2 and n+3, and one with strip n+4. The RDO truth objects are however split to document each individual strip, so there will be five associated RDO truth objects, each one containing the list of barcodes for each track, which deposited energy in the given strip, and the amount of energy deposited by each one of these tracks. For strip hits purely due to noise, the corresponding RDO truth object will obviously not refer the user back to any of the generated tracks.

2) The pixel RDOs are in one-to-one correspondence to each individual hit pixel and so are the corresponding pixel RDO truth objects, which are filled in the same way as the SCT RDO truth objects.

3) The TRT RDOs contain (or should contain) the full record of the low-threshold discriminator bit pattern in bins of 3.125 ns over three bunch crossings for straws with at least one time bin above threshold. The corresponding RDO truth objects should therefore contain the full list of tracks having contributed to setting the discriminator above threshold in one of the relevant time bins. I will try to check this before the meeting on 11/05.

Questions to the muons: 1) How and what is stored in the RDOs for each of the four technologies? 2) For the MDTs, it is my understanding that only the first hit to arrive on the wire will be recorded. Nevertheless, shouldn't the RDO truth object contain the full list of tracks which actually did deposit energy in the tube during the relevant time interval? What is the duration of the relevant time interval, the full drift time of about 500 ns or less? 3) For the RPCs, which information is stored in the RDOs? Single strips or sets of intersecting strips which actually triggered LVL1? What about the TGCs?

- From RDOs to PrepRawData objects, a decision is made to associate to each PrepRawData object one and only one track, which is the one having contributed the most energy to this object. PrepRawData are not saved to ESD. RIOonTrack objects are saved to ESD but have no truth associated to them.

- For each reconstructed track object, a simple sort of majority logic then selects the truth track which contributed the largest number of hits to the track. A pointer to this track and the fraction of hits it contributed are what the user can access from TrackTrack. If several TrackCollections are stored to ESD (e.g. iPatRec and xKalman), only one set has any truth information.

II) Requirements (as they have been since 1996, at least for the ID)

- for hits of any form (RDOs, PrepRawData, RIOonTrack), each hit should contain in some form the list of tracks which contributed to its creation. The amount of energy deposited is largely irrelevant since this can very rarely be of any use, especially with detectors using almost exclusively binary readout. The tracks should be stored at this stage as those which directly contributed to the hit, i.e. they can be primary or secondary

tracks

- for the track itself, the user may require access to the full list of associated hits and their truth information, stored according to the above requirements and also more directly to the truth track which generated most of its hits, along with the fraction of hits it actually contributed to. Here the truth track should be taken always as the original primary, or as close to that as one can track back along the parents. As examples, one can quote: a) an electron which lost energy repeatedly through a series of soft brem processes. Whether or not its curvature changed significantly, the track truth information should direct the user back to the primary electron. A similar use-case can be imagined for a muon undergoing catastrophic brem in the calorimeter etc... b) a charged pion undergoing an inleastic interaction in the ID material. Many of these survive the interaction and only their barcode allows one to easily differentiate between them and their "daughters". These "daughters" sometimes carry most of the initial momentum, sometimes not. Here also, a choice will have to be made for an optimal association to the primary truth track. c) an electron from a photon conversion in the beam pipe, which then underwent a few brem processes. Here the parent "primary" should be the relevant electron from the photon conversion and not the photon itself.

- for all the detailed studies of fake tracks, ghost tracks, spoilt hits, ambiguous hits etc... done for the ID TDR, the above information was used at various stages. Thanks to the ability to do this, the limitations of the b-tagging algorithms used at the time were traced back to often arise from reconstructed secondaries (e.g. from photon conversions in one of the first two pixel layers) together with some amount of confusion in the B-layer itself due to large clusters arising from multiple tracks in the dense core of high-pT jets.

Questions to the muons: 1) Would the above cover also all muon requirements? 2) Should the "primaries" be considered as the charged particles emerging from the calorimeter in the case of the muons? What about all the background hits when they originate from charged particles and actually create correlated RPC hits for example?

III) A proposal for implementation (hopefully valid for ID and muons)

1) Make sure the RDO truth objects are created for the elementary detector elements (see for example the case of the SCT described above) and that the full list of tracks contributing to the actual signal recorded in the detector element is stored, even if only e.g. the drift-time information of the first arrival is actually read out from the ROD as in the case of the MDTs

2) Associate to each PrepRawData object a multi-map of truth tracks (the list should be sufficient without any information about the deposited energy). Make sure noise is correctly treated (the absence of truth information should be a reliable way of labelling noise hits)

3) Associate to each RIOonTrack object the same information cloned from the corresponding PrepRawData object. In this way, all TrackCollections will be able to point back to the truth information of their hits and this information will be stored to ESD without any serious penalty in terms of space.

4) Based on past and current experience and on ongoing/future validation work with the truth related to hadronic interactions, define precisely which majority logic should be used to associate the most likely truth track to each reconstructed track object. A number of decisions will have to be taken at this stage and they may have to be revised as work and understanding progresses. As an initial proposal, one could start with the following: for all cases, loop through the truth map associated with each RIOonTrack object and count how many times each primary track contributes (either directly or as surviving interactions of different types in the ID material, with the barcode used as the discriminator) and add to these the secondary charged particles which originate from a neutral mother (photons, K0, Lambda, etc...).

Questions to the muons: 1) Would the above seem reasonable as a first step to be applied perhaps for validation to the ID and only a bit later to the muons? 2) Should track segments be implemented in the same way as tracks, i.e. through their associated hits on one side and an associated truth track based on a majority

logic on the other? 3) Do use-cases exist from past and current experience where such information will actually be required? It seems to me that such features will be useful to trace back e.g. the causes for spurious LVL1 triggers and/or low momentum fake tracks produced by the reco code in conditions of high background rates.

-- SolveigAlbrand - 06 Jun 2005

This topic: Sandbox > SuggestedDuringTheReview
Topic revision: r1 - 2005-06-06 - SolveigAlbrand