Analysis Tools / Fall MC Challenge

Dec. 6, 2012

H. Greenlee

General Topics

- MC samples.
- Algorithms.
- Other larsoft release issues.
- Auxilliary data (flux files).
- Tools.
- Documentation.

MC Samples

- I made a Google spreadsheet with a proposed list of samples (url too long to include here; ask me for an e-mail).
  - Worksheets: single, beam, other (incl. cosmics).
- Need to store all required generator fcl files in products/ubfcl in ubooneoffline repository.
- Each sample also needs a project xml file in products/ubxml.
  - ubxml is newly added product package since last meeting.

Single Particles

	A	B	C	D	E	F
1	Particle	Momentum (GeV)	Angular Dist.	Events	Event Size (Mb)	Sample Size (Gb)
2						
3	mu-	0.1-5.0	Forward (sigma=25 deg)	10000	2	20
4	mu-	0.1-2.0	Isotropic	10000	2	20
5	mu+	0.1-5.0	Forward (sigma=25 deg)	10000	2	20
6	mu+	0.1-2.0	Isotropic	10000	2	20
7	e-	0.1-5.0	Forward (sigma=25 deg)	10000	2	20
8	e-	0.1-2.0	Isotropic	10000	2	20
9	e+	0.1-5.0	Forward (sigma=25 deg)	10000	2	20
10	e+	0.1-2.0	Isotropic	10000	2	20
11	pi+	0.1-2.0	Isotropic	10000	2	20
12	pi-	0.1-2.0	Isotropic	10000	2	20
13	K+	0.1-2.0	Isotropic	10000	2	20
14	K-	0.1-2.0	Isotropic	10000	2	20
15	p	0.1-2.0	Isotropic	10000	2	20
16	gamma	0.1-2.0	Isotropic	10000	2	20
17	pi0	0.1-2.0	Isotropic	10000	2	20
18	n	0.1-2.0	Isotropic	10000	2	20
19	lambda	0.1-2.0	Isotropic	10000	2	20
20	K0S	0.1-2.0	Isotropic	10000	2	20
21	K0L	0.1-2.0	Isotropic	10000	2	20
22	nbar	0	Isotropic	10000	2	20
23						
24	Total			200000		400

Beam

	A	B	C	D	E	F
1	Generaator	Sample	Flux	Events	Event Size (Mb)	Sample Size (Gb)
2						
3	Genie	numu CC inclusive	BNB	10000	2	20
4	Genie	anumu CC inclusive	BNB	10000	2	20
5	Genie	numu NC inclusive	BNB	10000	2	20
6	Genie	anumu NC inclusive	BNB	10000	2	20
7	Genie	nue CC inclusive	BNB	10000	2	20
8	Genie	anue CC inclusive	BNB	10000	2	20
9						
10	Total			60000		120

Other

	A	B	C	D	E
1	Generaator	Sample	Events	Event Size (Mb)	Sample Size (Gb)
2					
3	CRY	Cosmic muons	10000	2	20
4		Supernova			
5		Nucleon decay			
6					
7	Total		10000		20

Algorithms

- Here is a list of candidate algorithms with my current assessment of readiness and/or usefulness.
  - Fuzzy clusters (no).
  - Bezier tracks (no).
  - Seeds (maybe).
  - Optical (yes, also needs to be added in sim).
  - Shower reco (no).
  - Space points (maybe).
  - Cheater reco (no, can't easily coexist with standard reco).
  - 3D Kalman Hit track (yes).

Larsoft Release Issues

- Modified Hit to allow backtracking if Wires are dropped (done).
  - Added Ptr<RawDigit> in Hit.
  - Wires should/will be dropped.
- Fix event display to work reasonably if Wires are dropped (not done, but generation can start before this is finished).
- Updated ProdSingle to generate “uniform” angular distribution flat in phase space (done).
- Remove debugging messages from SeedFinder (not done).
- Final algorithm updates (optical?).

Output Size Analysis for Genie Events

Output Size Including Wires

```
Branches of Events tree from all files:

   Total bytes  Zipped bytes   Comp.  Branch name
   -----------  ------------   -----  -----------
          3775           355   10.63  art::TriggerResults_TriggerResults__GenieGen.obj
          3771           352   10.71  art::TriggerResults_TriggerResults__Reco.obj
    2652872959      11790734  225.00  raw::RawDigits_daq__GenieGen.obj
      14539838        579825   25.08  recob::Clusterrecob::Hitvoidart::Assns_dbcluster__Reco.obj
       1071643        376874    2.84  recob::Clusters_dbcluster__Reco.obj
       6674050        378187   17.65  recob::Hitrecob::SpacePointvoidart::Assns_trackkalmanhit__Reco.obj
          2238           308    7.27  recob::Hitrecob::Trackvoidart::Assns_trackkalmanhit__Reco.obj
      37703338      18122977    2.08  recob::Hits_ffthit__Reco.obj
       2766955         98494   28.09  recob::SpacePointrecob::Trackvoidart::Assns_trackkalmanhit__Reco.obj
       3423163       1336182    2.56  recob::SpacePoints_trackkalmanhit__Reco.obj
      13331661       5715436    2.33  recob::Tracks_trackkalmanhit__Reco.obj
    5301575607    1217161200    4.36  recob::Wires_caldata__Reco.obj
     137485246     107151244    1.28  sim::SimChannels_largeant__GenieGen.obj
          1585           667    2.38  sim::SimPhotonss_largeant__GenieGen.obj
         34946         23461    1.49  simb::GTruths_generator__GenieGen.obj
          5640           375   15.04  simb::GTruthsimb::MCTruthvoidart::Assns_generator__GenieGen.obj
         51941         19337    2.69  simb::MCFluxs_generator__GenieGen.obj
          5640           375   15.04  simb::MCFluxsimb::MCTruthvoidart::Assns_generator__GenieGen.obj
      24908710      10939366    2.28  simb::MCParticles_largeant__GenieGen.obj
        656592         26006   25.25  simb::MCParticlesimb::MCTruthvoidart::Assns_largeant__GenieGen.obj
        308308         71443    4.32  simb::MCTruths_generator__GenieGen.obj
    8197427606    1373793198    5.97  All branches

        50 events.
163.95 Mb average size per event.
 27.48 Mb average zipped size per event.
```

Output Size Analysis for Genie Events

Output Size Not Including Wires

Branches of Events tree from all files:

Total bytes	Zipped bytes	Comp.	Branch name
3775	355	10.63	art::TriggerResults_TriggerResults__GenieGen.obj
3771	352	10.71	art::TriggerResults_TriggerResults__Reco.obj
2652872959	11790734	225.00	raw::RawDigits_daq__GenieGen.obj
14539838	579825	25.08	recob::Clusterrecob::Hitvoidart::Assns_dbcluster__Reco.obj
1071643	376874	2.84	recob::Clusters_dbcluster__Reco.obj
6674050	378187	17.65	recob::Hitrecob::SpacePointvoidart::Assns_trackkalmanhit__Reco.obj
2238	308	7.27	recob::Hitrecob::Trackvoidart::Assns_trackkalmanhit__Reco.obj
37703338	18122977	2.08	recob::Hits_ffthit__Reco.obj
2766955	98494	28.09	recob::SpacePointrecob::Trackvoidart::Assns_trackkalmanhit__Reco.obj
3423163	1336182	2.56	recob::SpacePoints_trackkalmanhit__Reco.obj
13331661	5715436	2.33	recob::Tracks_trackkalmanhit__Reco.obj
137485246	107151244	1.28	sim::SimChannels_largeant__GenieGen.obj
1585	667	2.38	sim::SimPhotonss_largeant__GenieGen.obj
34946	23461	1.49	simb::GTruths_generator__GenieGen.obj
5640	375	15.04	simb::GTruthsimb::MCTruthvoidart::Assns_generator__GenieGen.obj
51941	19337	2.69	simb::MCFluxs_generator__GenieGen.obj
5640	375	15.04	simb::MCFluxsimb::MCTruthvoidart::Assns_generator__GenieGen.obj
24908710	10939366	2.28	simb::MCParticles_largeant__GenieGen.obj
656592	26006	25.25	simb::MCParticlesimb::MCTruthvoidart::Assns_largeant__GenieGen.obj
308308	71443	4.32	simb::MCTruths_generator__GenieGen.obj
2895851999	156631998	18.49	All branches

50 events.
57.92 Mb average size per event.
3.13 Mb average zipped size per event.

Flux Files and Auxilliary Data

- Identify and catalog existing flux files.

Tools

- Updated project.py to support merging/splitting (probably don't need right away).
- Added ubxml uboone product.
- Sam (not ready).
- Uboonepro account.
  - Login access OK.
  - Grid proxy issues solved (I think).

Documentation

- Web-browsable data catalog.
  - Not started. Not yet sure what is the best way to proceed.
  - Data generation doesn't need to wait for this.