Configuration for DNNROI training truth and fodder

Brett Viren

February 5, 2026

Topics

- Personal LLM-assisted s/w dev workflow
- Use and status of `dnnroi-training.jsonnet`
- Next steps

Beads for structured prompting of Claude

```
$ bd create --title 'Fix the thing'
--> wct-xyz
$ bd edit wct-xyz # write a few paragraphs of prompt
$ bd create --title 'Fix the other thing'
--> wct-abc
$ bd deb add wct-abc wct-xyz # add dependency
$ bd list
$ bd show wct-abc

$ claude
```

> *Read beads issue wct-xyz and address it by following the instructions in the the description. Ask questions if anything is uncertain. Update the issue with a summary of what you have done.*

Claude also learned to provide `doctest_*.cxx` that covers the work.

Time usage

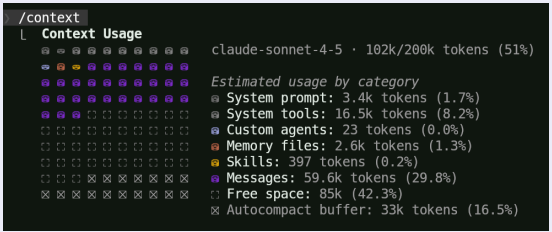
My time

heavy	writing detailed design descriptions in the Beads issues.
medium	checking resulting code and tests, running and fixing.
easy	monitoring Claude and (rare) countermanding a command.

Claude's time

- I used Sonnet 4.5. Opus would be even better (and more expensive).
- No `/usage` with Claude+Azure. From Azure, adding HDF5 **cost $10**.

Based on input token count, should more like **$1** (Sonnet: $3/M in, $15/M out)



Non Claude development areas

I still do not understand how I can use Claude for developing **novel, structural** Jsonnet because I discover what I want as I go and "right" is very subject to human "taste".

But, I expect that given current `subgraphs.jsonnet` it is possible to perform tasks within the envelope of known stuff.

Hmm… **let's try!**

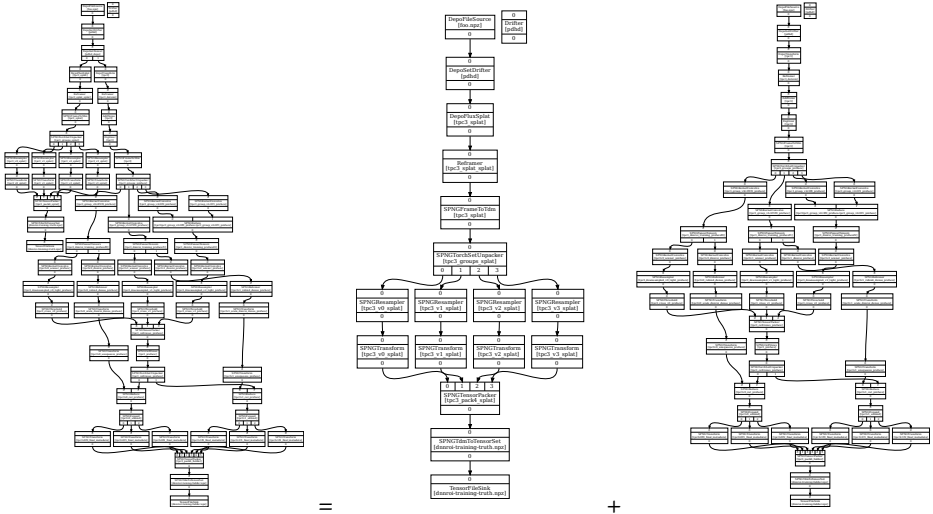
Claude tries Jsonnet

```
$ claude
```

> *Read spng/cfg/spng/subgraphs.jsonnet and make me two main Jsonnet files both similar to dnnroi-training.jsonnet. One is called dnnroi-training-truth.jsonnet and the other dnnroi-traning-fodder.jsonnet. Each should make a graph corresponding to each of these subgraphs made by dnnroi-training.jsonnet. Ask questions if you have uncertainties about what is being requested and how to perform the task.*

- Claude asks some reasonable and multi-choice questions with the default answers guessed correctly.
- Resulting Jsonnets looks decent and compile on first try.

How do they look?



Okay, enough messing around with Claude.

dnnroi-training.jsonnet

Many command line parameters!

input	The name of a file in WCT “depo file” format, usually .npz.
outpat	The output file name pattern with format variables.
schema	The output file schema (“tensor” or “frame”)
detname	The name of a supported detector, default “pdhd”.
engine	The name of the graph execution engine, default Pgrapher or TbbFlow.
device	The name of the device for SPNG nodes, default “cpu” or “gpu”, “gpu1”, etc.
rebin	The downsample factor along the time dimension.
scale	The amount to scale down the dense arrays (splat and looseLF)
tpcid	The TPC ID number.
dump	The list of intermediates to dump to file.
verbosity	The verbosity level for additional logging.

But mostly just need

```
$ wire-cell -c spng/cfg/spng/dnnroi-training.jsonnet \
            -A input=test/data/muon-depos.npz
```

Automatic output file format

HDF5 and NPZ:

```
$ wire-cell [...] -A outpat='dnnroi-training-%(tier)s.hdf
$ wire-cell [...] -A outpat='dnnroi-training-%(tier)s.npz
```

Current HDF5 schema

```
$ h5ls -fr dnnroi-training-fodder.hdf
/                        Group
/frames                  Group
/frames/0                Group
/frames/0/tags           Group
/frames/0/tags/fodder    Group
/frames/0/tags/fodder/view Group
/frames/0/tags/fodder/view/0 Group
/frames/0/tags/fodder/view/0/feature Group
/frames/0/tags/fodder/view/0/feature/dense Group
/frames/0/tags/fodder/view/0/feature/dense/traces Dataset {800, 1500}
/frames/0/tags/fodder/view/0/feature/mp2 Group
/frames/0/tags/fodder/view/0/feature/mp2/traces Dataset {800, 1500}
/frames/0/tags/fodder/view/0/feature/mp3 Group
/frames/0/tags/fodder/view/0/feature/mp3/traces Dataset {800, 1500}
/frames/0/tags/fodder/view/1 Group
/frames/0/tags/fodder/view/1/feature Group
/frames/0/tags/fodder/view/1/feature/dense Group
/frames/0/tags/fodder/view/1/feature/dense/traces Dataset {800, 1500}
/frames/0/tags/fodder/view/1/feature/mp2 Group
/frames/0/tags/fodder/view/1/feature/mp2/traces Dataset {800, 1500}
/frames/0/tags/fodder/view/1/feature/mp3 Group
/frames/0/tags/fodder/view/1/feature/mp3/traces Dataset {800, 1500}
/tensorsets              Group
/tensorsets/0            Group
/tensorsets/0/0          Dataset, same as /frames/0/tags/fodder/view/0/feature/dense/traces
/tensorsets/0/1          Dataset, same as /frames/0/tags/fodder/view/0/feature/mp2/traces
/tensorsets/0/2          Dataset, same as /frames/0/tags/fodder/view/0/feature/mp3/traces
/tensorsets/0/3          Dataset, same as /frames/0/tags/fodder/view/1/feature/dense/traces
/tensorsets/0/4          Dataset, same as /frames/0/tags/fodder/view/1/feature/mp2/traces
/tensorsets/0/5          Dataset, same as /frames/0/tags/fodder/view/1/feature/mp3/traces
```

HDF5 Serialization rules

- Write `ITensorSet` to `/tensorsets/{ident}`.
- Write `ITensor` under `/tensorsets/{ident}/{index}`.
- Use `"datapath"` metadata to make **HDF links**

The `/tensorsets/` path define container and the links provide more meaningful paths.

```
/frames/0/tags/fodder/view/0/feature/dense/traces -> /tensorsets/0/0
/frames/0/tags/fodder/view/0/feature/mp2/traces   -> /tensorsets/0/1
/frames/0/tags/fodder/view/0/feature/mp3/traces   -> /tensorsets/0/2
```

Is the current HD5 schema sufficient for `wirecell-dnn`

- Content is what DNNROI needs, JakeNet ignores mp2/mp3 uses “dense”.
- At least need to change `wirecell-dnn` config for new regexs

Next steps

Start a `spng/test/spng` Snakemake script

- Use trail blazed by Jake's `training_input/Snakefile`
  - Container based running of art/LArSoft
- Add rules:
  - LArG4 + `dnnroi-training.jsonnet` for training sample
  - Run `wcpy dnn train` to make DNNROI network
  - LArG4 + `dnnroi-inference.jsonnet` for SP on cosmics
  - SPDIR + `dnnroi-inference.jsonnet` for plots