Accessing and using weather data in OCaml

Hez Carty - OCaml 2013
MDA Information Systems LLC
Weather and OCaml

- OCaml - why and where?
- Library bindings
- Highs and lows
Why OCaml

● The usual reasons
  ○ Functional (when you want to be)
  ○ Type safe and expressive
  ○ Native code (fast), bytecode (REPL)
  ○ Predictable

● Relatively simple FFI
  ○ Bigarrays, C-friendly float arrays

● Reasonable selection of native and bound libraries
What OCaml

● Weather data retrieval and validation
● Services
  ○ HTTP, zeromq
  ○ Task/workflow management
● Data analysis and reduction
  ○ Teleconnections (PCA, plotting)
  ○ Precipitation probability model (Monte Carlo simulation with a first order Markov chain)
  ○ General data processing and preparation
    ■ Climatology
    ■ Data extraction/insertion
Raw bits and bytes
HDF4 and GRIB

- Binary data formats used in Earth sciences
- C libraries, lots of analysis tools
- Multi-dimensional data
- HDF4 bindings
  - Bigarrays wrapped in a variant + converters
  - Mixture of hand-written and camlidl
  - Low level and OCaml-friendly interfaces
- GRIB bindings
  - ints, floats, float arrays
  - Bindings are entirely hand-written
  - Only exposes an OCaml-friendly interface
Derived data - Teleconnections
Challenges

● Standard issues when interfacing with C
  ○ Type mismatches between C and OCaml, void *
  ○ Share vs copy?
● Balancing elegance, purity with potential speed + memory savings of mutation
● Development environments other than emacs and vim
● Accessibility and exposure to people outside of the community
Successes

- HDF4 and GRIB API bindings are/have been used to process terabytes of data per day
- Bigarrays with mmap + the FFI makes getting up and running quick and easy
- opam with local repositories
  - GODI, odb/oasis-db before that
- utop, merlin, ocamlbrowser
- ocp-indent
- Lwt, Batteries, cohttp, GSL, atdgen
Thanks!
Other related bindings

● UDUNITS
  ○ Unit conversion
● PROJ.4
  ○ Coordinate projections
● PLplot
  ○ Various kinds of plots and maps
● Proper/official opam packages coming
● Until then:
  ○ http://github.com/hcarty/
  ○ http://0ok.org/ocaml/repo/ - opam repository