Retrospect Implementation Notes

* Class rewriting
  + JVM doesn’t allow manually defining class in java.\* package, so have to rename at least some classes
* Code size
  + Need to guarantee code doesn’t get too much bigger!
* Object initialization
  + Native objects
    - Generate no-argument constructor
    - Technically works – VM doesn’t verify final fields are assigned
    - Not ideal because VM might be free to cache values, ignore assignments, etc.
    - Could also use FieldMapMirrors instead – may be faster than reflection-based NativeObjectMirrors
    - Should refactor mirror interface to use constant pool offsets instead of field names
      * Can dynamically generate native mirrors that use a switch to select the right field
    - Note: two different kinds of “native” objects
      * Instantiated inside MirageClassLoader classes – all fields lifted to mirages already
      * Instantiated outside – all fields have to be lifted as they are accessed
    - Direct datatype implementation may still be fastest since all fields must be of type ObjectMirror?
      * Maybe not, can still separate primitives from references
  + Mirage objects
    - Add extra mirror field to constructors, pass up super chain
      * Doesn’t handle reflection (i.e. Class#newInstance)
      * In fact, messes up reflection entirely since the parameter types are wrong
        + Need to advice reflection calls to hide changes in general – aligns with principle of mirror-based architectures
      * How to distinguish calls to new (where mirror should be instantiated) from calls to super (where mirror should be passed along?)
        + Both are just INVOKESPECIAL
        + Do dataflow analysis that respects difference between uninitialized and uninitialized this
        + Mirror instantiated on caller side
      * Reflective construction (where mirror is already available) uses constructor that takes an Object
        + Must have different signature to distinguish from when the original class has a nullary constructor, to which we add the extra mirror argument
    - Instead, two ObjectMirage constructors (NO)
      * Nullary for new statements – creates native mirror inline
      * One that takes a mirror for calls to ObjectMirage.make
      * Problem
        + In Java, anonymous inner class constructors will set fields before calling super
        + Bytecode in general can do this whenever it wants
* Static fields
  + ClassMirror interface for getting fields
  + Single ClassMirrorLoader instance per ClassLoader
  + Problem – can’t ClassLoader.loadClass() on non-public classes, even though it’s easy for many classes to get a hold of the Class instance by calling getClass() on an instance
* ClassLoading:
  + MirageClassLoader
    - Give it two class loaders for delegation
      * One for mirror implementations
      * One for loading original code (for transforming to mirage classes)
    - Desirable to catch setup errors where bytecode provided by class loader doesn’t match original bytecode
      * Any missing/incorrectly typed fields
      * Any incorrect *serialVersionUID* fields?
  + Mirrors turn out to be useful for bytecode verification/analysis
    - Type checking and analysis of mirror tree instead of real classes avoids class circularity errors
* Arrays
  + Each unique type generates a stub Java class that inherits ObjectArrayMirror
    - No instance methods no not necessary for the same reason as user classes, but needed for method overloading
  + Big issue: AALOAD/STORE opcode doesn’t specify expected element type the way method invocations do
    - Need to analyze bytecode to infer it!
    - Need bytecode version 1.6 or later - frame information required
      * No good, JRE only 1.5 even in later versions
* Special cases
  + All classes must extend Object, which declares toString()
    - Therefore all toString() methods must return actual java.lang.String instances
  + Object identity
    - How to define ==?
      * Could define ObjectMirror.sameObject(Object other)
      * Could allow mirages to be different, even if calling ObjectMirage.make() on the same mirror object twice
    - How to define hashCode()?
      * Supposed to be consistent across a single VM execution
    - How to define equals()?
    - Current approach: mirror identity implies mirage identity, lots of lazy maps to ensure same instances
  + Native methods
    - Enumerate replacements
      * Use AspectJ binaries even at this point (around execution)?
    - Automatically link all extra methods on ClassMirror#getNativeStubsClass according to idiom
      * How to deal with overloading? Stubs will all take ObjectMirrors instead of specific classes
    - Any missing methods are stubbed with a method that throws an exception
    - Need to easily support native methods outside JVM/JRE
  + Reflection
    - Should intercept methods like Class.getFields()/newInstance()/getName() etc. to interface with mirrors instead
  + Enums
    - Just remove the flag – no need?
  + Soft/weak/etc reference
    - ???
* Optimization ideas
  + Cache consecutive field reads on “this” in same method?
    - No, not valid in light of concurrent writes
    - Maybe in specific well-understood/constrained cases
    - Or with clearer understanding of Java memory model
* Memory model
  + Mutable mirrors must handle concurrent access correctly
    - FieldMapMirror needs to be backed by ConcurrentMap
    - DirectArrayMirror needs to be backed by (array of AtomicReferences?)
    - Native methods like Unsafe.compareAndSwapInt() should call similar methods on FieldMirror
      * (in concert with Unsafe.objectFieldOffset)
    - Probably want “threadsafe” flag to save synchronization costs in common case