INPUT:

$code[]$ - An array of disassembly output

OUTPUT:

IS - Initial set of embedded data

AS - The set of seed addresses for embedded data

RS - The set of reference sites to embedded data

```
1:  procedure InitialSetCollection
2:      IS = {}
3:      AS = {}
4:      RS = {}
5:      for each (ldr-literal addr) ∈ code[] at curr do
6:          size = MemLoadSize(ldr)
7:          IS = IS ∪ {addr, addr+1, ..., addr+size-1}
8:          RS = RS ∪ {curr}
9:      end for
10:     for each (adr xn, addr) ∈ code[] at curr do
11:         escaped, depInsts = ForwardSlicing (xn)
12:         unbounded = False
13:         for each inst ∈ depInsts do
14:             if inst is MemoryLoad then
15:                 RS = RS ∪ {curr}
16:                 addr_expr = MemLoadAddrExpr(inst)
17:                 if IsBounded(addr_expr) then
18:                     IS = IS ∪ {AddrRange(addr_expr)}
19:                 else
20:                     unbounded = True
21:                 end if
22:             end if
23:         end for
24:         if escaped or unbounded then
25:             AS = AS ∪ {addr}
26:         end if
27:     end for
28: end procedure
```