Taverna → OPM basic mapping

- ok to represent specific lineage graphs
  - eg results of queries

- processor I/O var names also lost (“P₂V₁₁”)

- fine as long as granularity is constant
  - i.e. no collections

PR = “producer role”
CR = “consumer role”
**Granularity issues**

**PV**: $s = [a, b, c]$

**PV**: $s = [a', b', c']$

$L = [a, b, c]$  
$L' = [a', b', c']$

need to capture:  
$a \in L$, $b \in L$, $c \in L$  
$a' \in L'$, $b' \in L'$, $c' \in L'$

this may lose information

is this valid OPM?:

and also:

$P_1$
Annotations in OPM

\( PV_i: I(s) = [a, b, c] \)

\( PV_o: I(s) = [x, y] \)

\( PV_o: I(s) = [a', b', c'] \)

only useful annotation:
\( P \) is **index-preserving**:
\( PV_o[i] = PV_i[i] \)
\( \text{lineage}(PV_o[i]) = PV_i[i] \)

The annotation is used at query time to retain fine granularity would it be useful to push it into the provenance graph? how about using the annotation to deal with the granularity issue in the previous slide

\([a, b, c]\) \( \xrightarrow{\text{wgb}(PR)} \) \( P_1 \) \( \xrightarrow{\text{used}(CR)} \) \( [a', b', c'] \)

\( P_1 \text{ is index-preserving} \)

\( \text{hasAnnotation} \)