

Coherent Cluster

- For a 2x2 matrix containing two objects {x,y} and two attributes {a,b}

- $D = (X_a - Y_a) - (X_b - Y_b)$

- (to measure the shift in the data)

	Attribute a	Attribute b
Object x	Xa	Xb
Object y	Ya	Yb

- If $D \leq \delta$, then this 2x2 matrix is a δ -coherent cluster
 - An $m \times n$ matrix X is a δ -coherent cluster if every 2x2 submatrix of X is a δ -coherent cluster.
 - A δ -coherent cluster is a maximum δ -coherent cluster if it is not a submatrix of any other δ -coherent cluster.

Coherent Cluster

- **Objective:** given a data matrix and a threshold δ , find all maximum δ -coherent clusters.
- **Input:** matrix M , threshold δ
- **Algorithm:**
 1. Compute the **maximum coherent attribute sets** for each pair of objects
 2. Two-way pruning
 3. Construct the lexicographical tree
 4. Use **post-order traverse** the tree to find all maximum coherent clusters

Coherent Cluster

- Example:

	a_0	a_1	a_2	a_3
o_0	1	4	2	5
o_1	2	5	5	8
o_2	3	6	5	7
o_3	4	20	7	2
o_4	30	7	6	6

$\delta = 1$

Coherent Cluster

- Example:

	a_0	a_1	a_2	a_3
o_0	1	4	2	5
o_1	2	5	5	8
o_2	3	6	5	7
o_3	4	20	7	2
o_4	30	7	6	6

$\delta = 1$

1. Compute the **maximum coherent attribute sets** for each pair of objects

– (O_0, O_1)

	a_0	a_1	a_2	a_3
o_0	1	4	2	5
o_1	2	5	5	8
$D(O_0, O_1)$	-1	-1	-3	-3

Sort by D



	a_2	a_3	a_0	a_1
o_0	2	5	1	4
o_1	5	8	2	5
$D(O_0, O_1)$	-3	-3	-1	-1

$(o_0, o_1) : \{a_0, a_1\}, \{a_2, a_3\}$

– (O_0, O_3)

	a_0	a_1	a_2	a_3
o_0	1	4	2	5
o_3	4	20	7	2
$D(O_0, O_3)$	-3	-16	-5	3

Sort by D



	a_1	a_2	a_0	a_3
o_0	4	2	1	5
o_3	20	7	4	2
$D(O_0, O_3)$	-16	-5	-3	3

$(o_0, o_3) : \text{empty}$

Two Way Pruning

	a0	a1	a2
o0	1	4	2
o1	2	5	5
o2	3	6	5
o3	4	200	7
o4	300	7	6

delta=1 nc =3 nr = 3

(o0,o2) \rightarrow (a0,a1,a2)
 (o1,o2) \rightarrow (a0,a1,a2)

~~(o0,o2) \rightarrow (a0,a1,a2)~~
~~(o1,o2) \rightarrow (a0,a1,a2)~~

MCAS

(a0,a1) \rightarrow (o0,o1,o2)
 (a0,a2) \rightarrow (o1,o2,o3)
 (a1,a2) \rightarrow (o1,o2,o4)
 (a1,a2) \rightarrow (o0,o2,o4)

~~(a0,a1) \rightarrow (o0,o1,o2)~~
~~(a0,a2) \rightarrow (o1,o2,o3)~~
~~(a1,a2) \rightarrow (o1,o2,o4)~~
~~(a1,a2) \rightarrow (o0,o2,o4)~~

MCOS

Coherent Cluster

- After step 1:

$(o_0, o_1) : \{a_0, a_1\}, \{a_2, a_3\}$

$(o_0, o_2) : \{a_0, a_1, a_2, a_3\}$

$(o_0, o_4) : \{a_1, a_2\}$

$(o_1, o_2) : \{a_0, a_1, a_2\}, \{a_2, a_3\}$

$(o_1, o_3) : \{a_0, a_2\}$

$(o_1, o_4) : \{a_1, a_2\}$

$(o_2, o_3) : \{a_0, a_2\}$

$(o_2, o_4) : \{a_1, a_2\}$

2. Two-way pruning

- Apply for constraints
 - Ex: min number of object
 - Ex: min number of attribute
- Remove the coherent attribute/object sets that do not satisfy our constraints
- In this example, this step is omitted

Coherent Cluster

- After step 1

$(o_0, o_1) : \{a_0, a_1\}, \{a_2, a_3\}$
 $(o_0, o_2) : \{a_0, a_1, a_2, a_3\}$
 $(o_0, o_4) : \{a_1, a_2\}$
 $(o_1, o_2) : \{a_0, a_1, a_2\}, \{a_2, a_3\}$
 $(o_1, o_3) : \{a_0, a_2\}$
 $(o_1, o_4) : \{a_1, a_2\}$
 $(o_2, o_3) : \{a_0, a_2\}$
 $(o_2, o_4) : \{a_1, a_2\}$

3. Construct the lexicographical tree

- Rewrite the maximal coherent attribute sets:

$\{a_0, a_1\} : (o_0, o_1)$
 $\{a_0, a_2\} : (o_1, o_3), (o_2, o_3)$
 $\{a_1, a_2\} : (o_0, o_4), (o_1, o_4), (o_2, o_4)$
 $\{a_2, a_3\} : (o_0, o_1), (o_1, o_2)$
 $\{a_0, a_1, a_2\} : (o_1, o_2)$
 $\{a_0, a_1, a_2, a_3\} : (o_0, o_2)$

Coherent Cluster

3. Construct the lexicographical tree

$$\{a_0, a_1\} : (o_0, o_1)$$

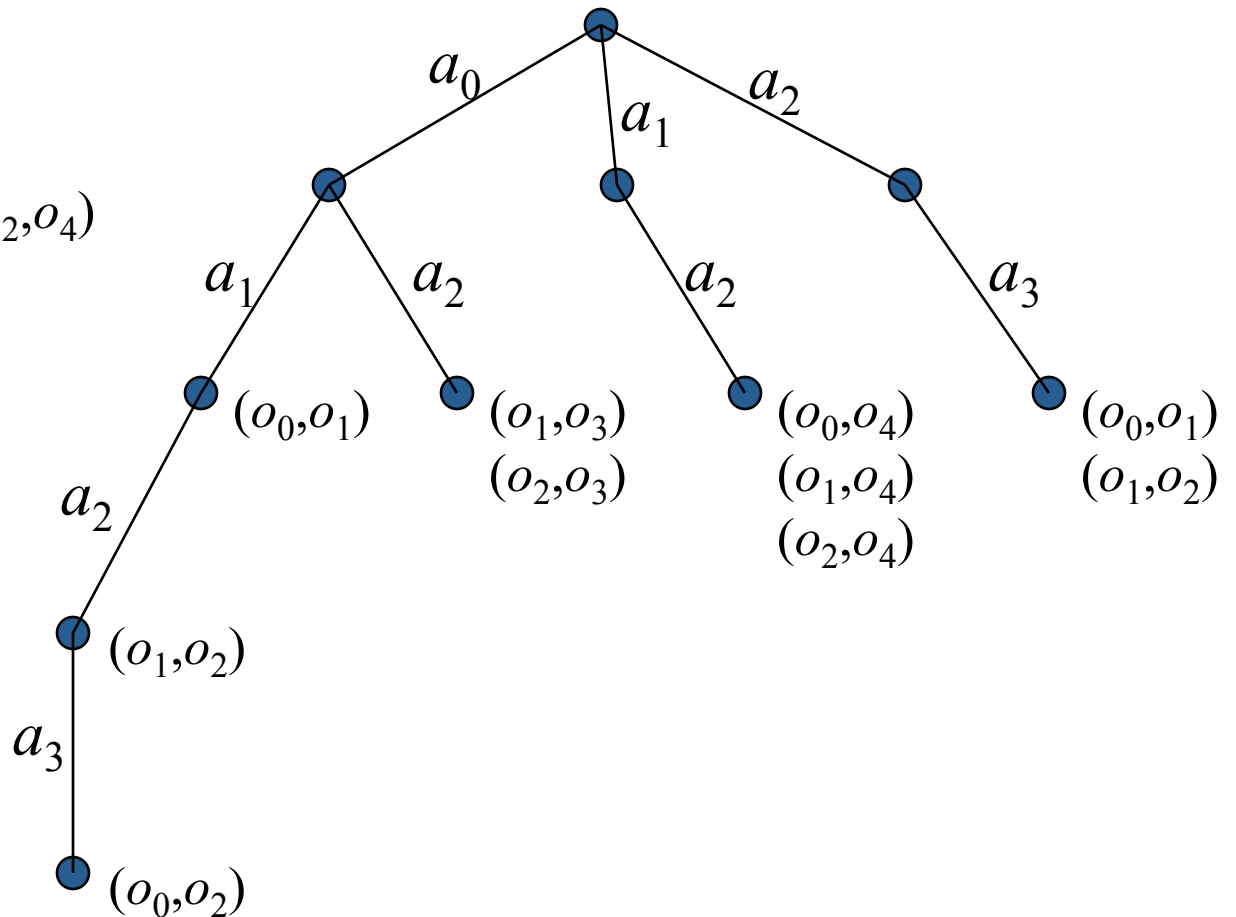
$$\{a_0, a_2\} : (o_1, o_3), (o_2, o_3)$$

$$\{a_1, a_2\} : (o_0, o_4), (o_1, o_4), (o_2, o_4)$$

$$\{a_2, a_3\} : (o_0, o_1), (o_1, o_2)$$

$$\{a_0, a_1, a_2\} : (o_1, o_2)$$

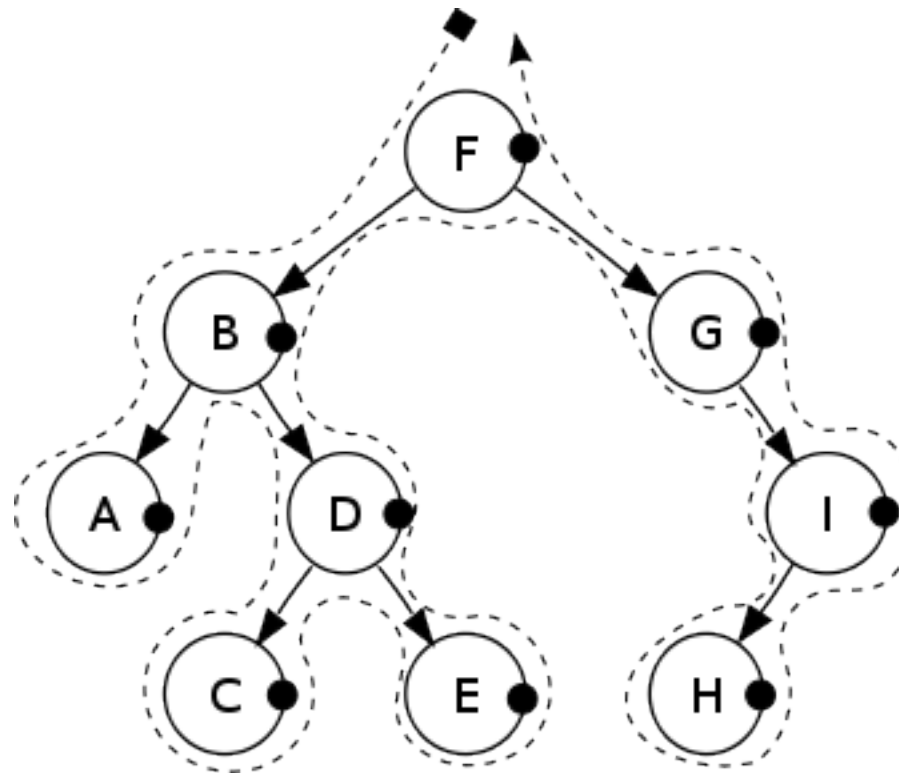
$$\{a_0, a_1, a_2, a_3\} : (o_0, o_2)$$



Coherent Cluster

4. Use **post-order traverse** the tree to find all maximum coherent clusters

Post-order Traversal: A – C – E – D – B – H – I – G – F



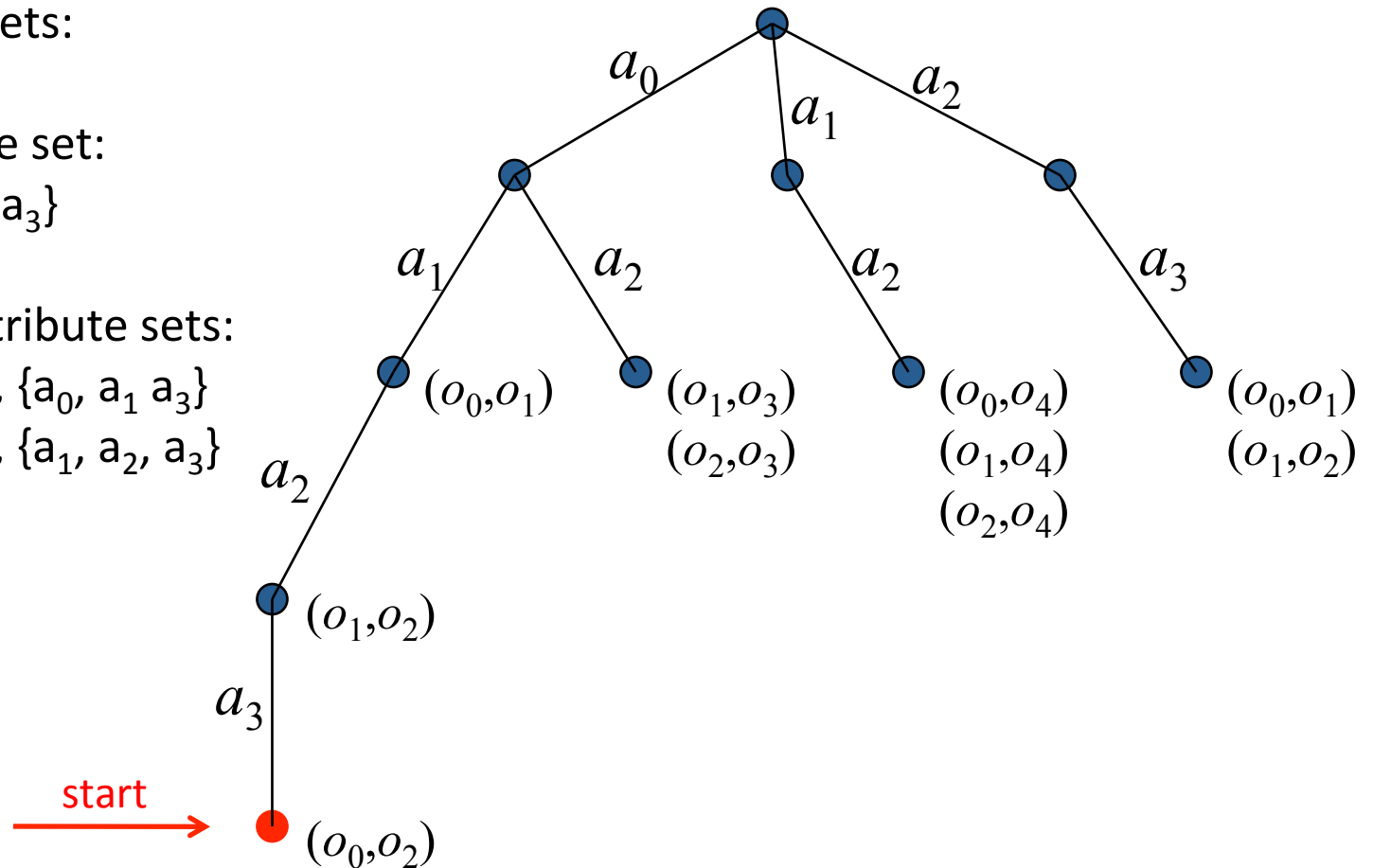
Coherent Cluster

4. Use **post-order traverse** the tree to find all maximum coherent clusters
 - At each node
 - Represented by a coherent attribute set and its coherent object sets
 - **Copying:**
 - » Let k = the size of the coherent attribute set
 - » Copy the object sets to all nodes that represent the subsets of the coherent attribute set with $(k-1)$ size
 - » Purpose – to facilitate the recovery of the maximum coherent clusters
 - **Report:** Report the maximum coherent clusters

Coherent Cluster

4. Use **post-order traverse** the tree to find all maximum coherent clusters

- Coherent object sets:
 (O_0, O_2)
- Coherent attribute set:
 $\{a_0, a_1, a_2, a_3\}$
- (K-1) Coherent attribute sets:
 $\{a_0, a_1, a_2\}, \{a_0, a_1, a_3\}$
 $\{a_0, a_2, a_3\}, \{a_1, a_2, a_3\}$



Coherent Cluster

4. Use **post-order traverse** the tree to find all maximum coherent clusters

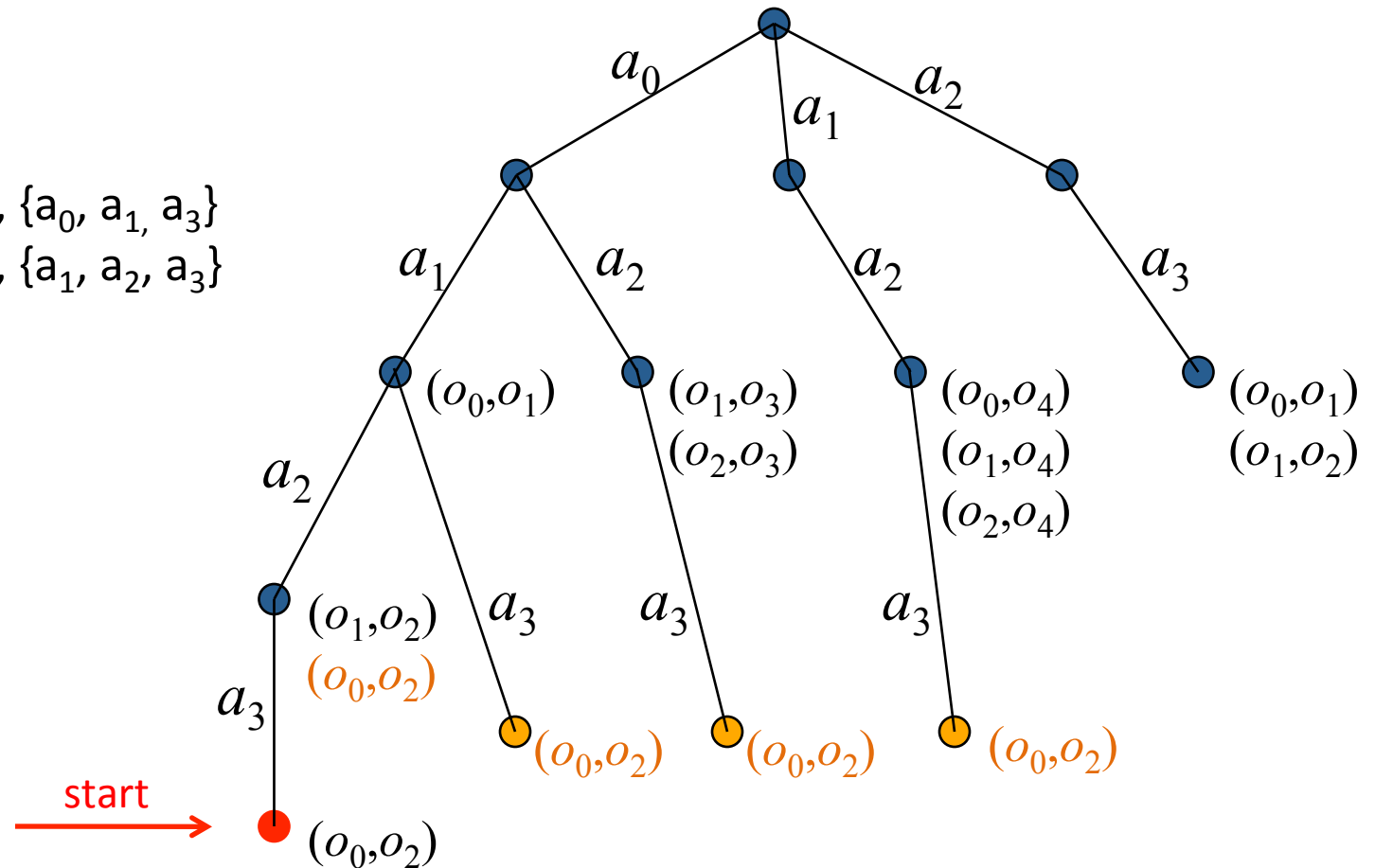
- **Copying**

(O_0, O_2)

to

$\{a_0, a_1, a_2\}, \{a_0, a_1, a_3\}$

$\{a_0, a_2, a_3\}, \{a_1, a_2, a_3\}$

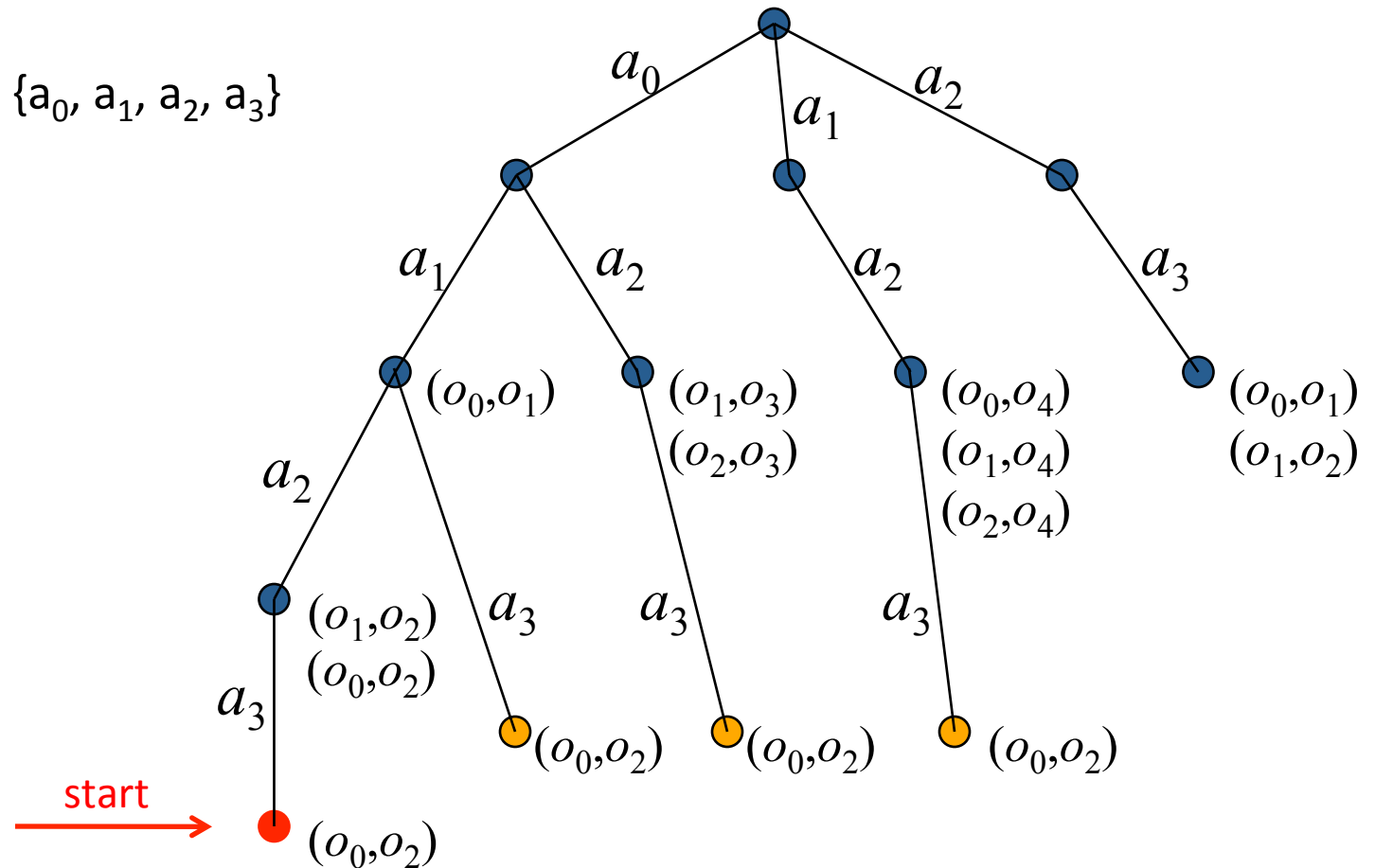


Coherent Cluster

4. Use **post-order traverse** the tree to find all maximum coherent clusters

- **Report:**

$$\{O_0, O_2\} \times \{a_0, a_1, a_2, a_3\}$$



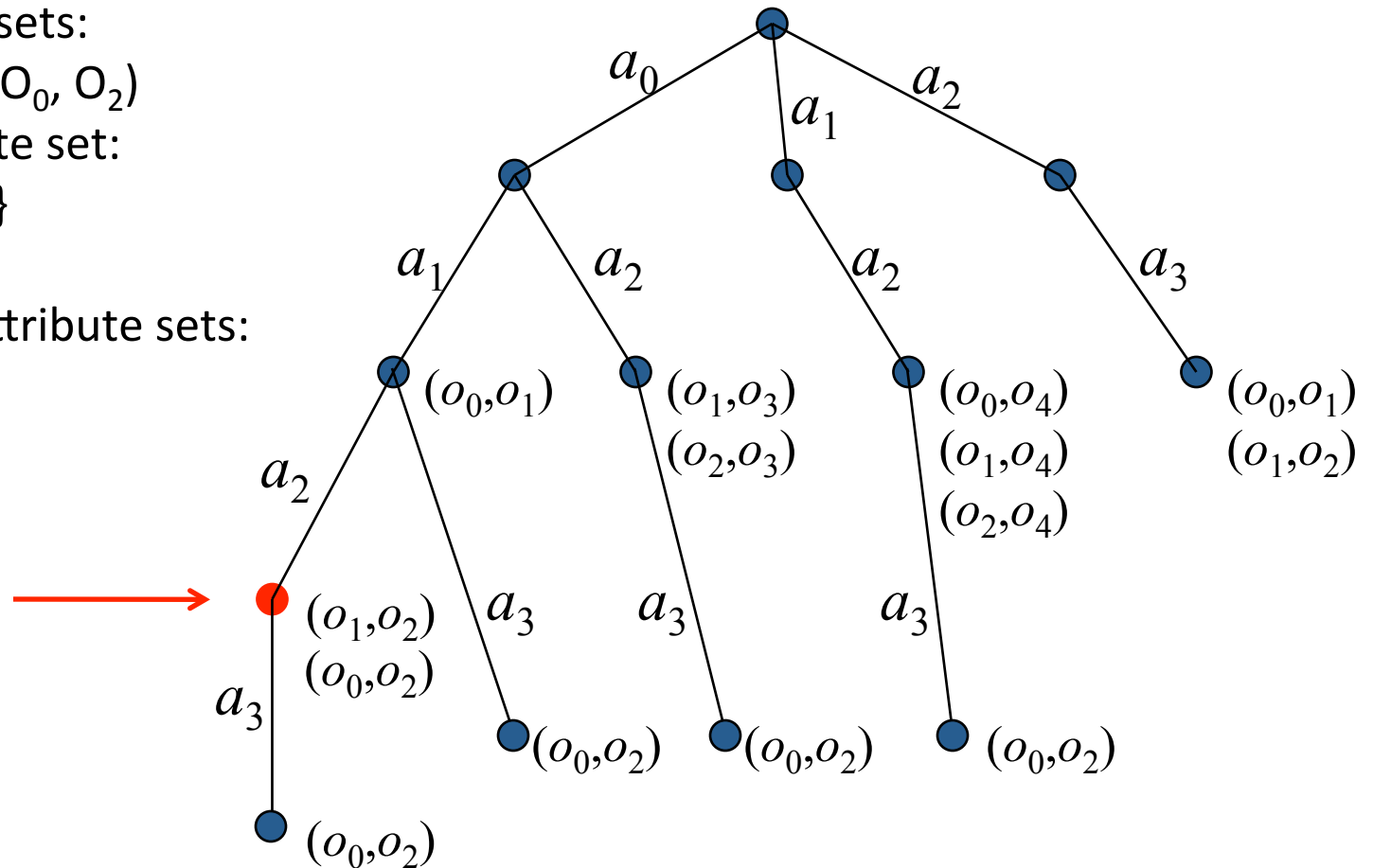
Coherent Cluster

4. Use **post-order traverse** the tree to find all maximum coherent clusters

- Coherent object sets:
 $(O_1, O_2), (O_0, O_2)$

- Coherent attribute set:
 $\{a_0, a_1, a_2\}$

- (K-1) Coherent attribute sets:
 - $\{a_0, a_1\}$,
 - $\{a_0, a_2\}$,
 - $\{a_1, a_2\}$



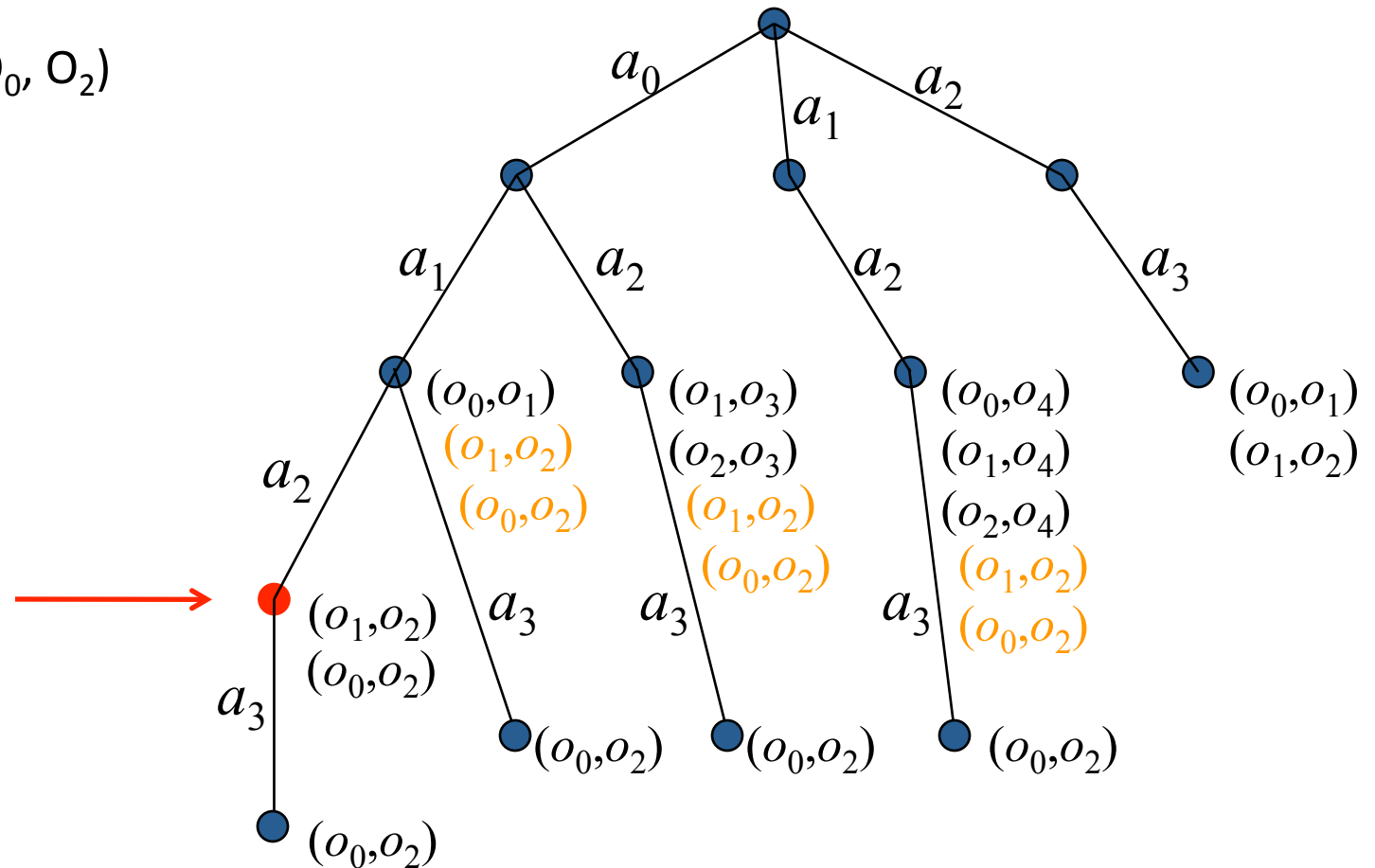
Coherent Cluster

4. Use **post-order traverse** the tree to find all maximum coherent clusters

- **Copying**

$$(O_1, O_2), (O_0, O_2)$$

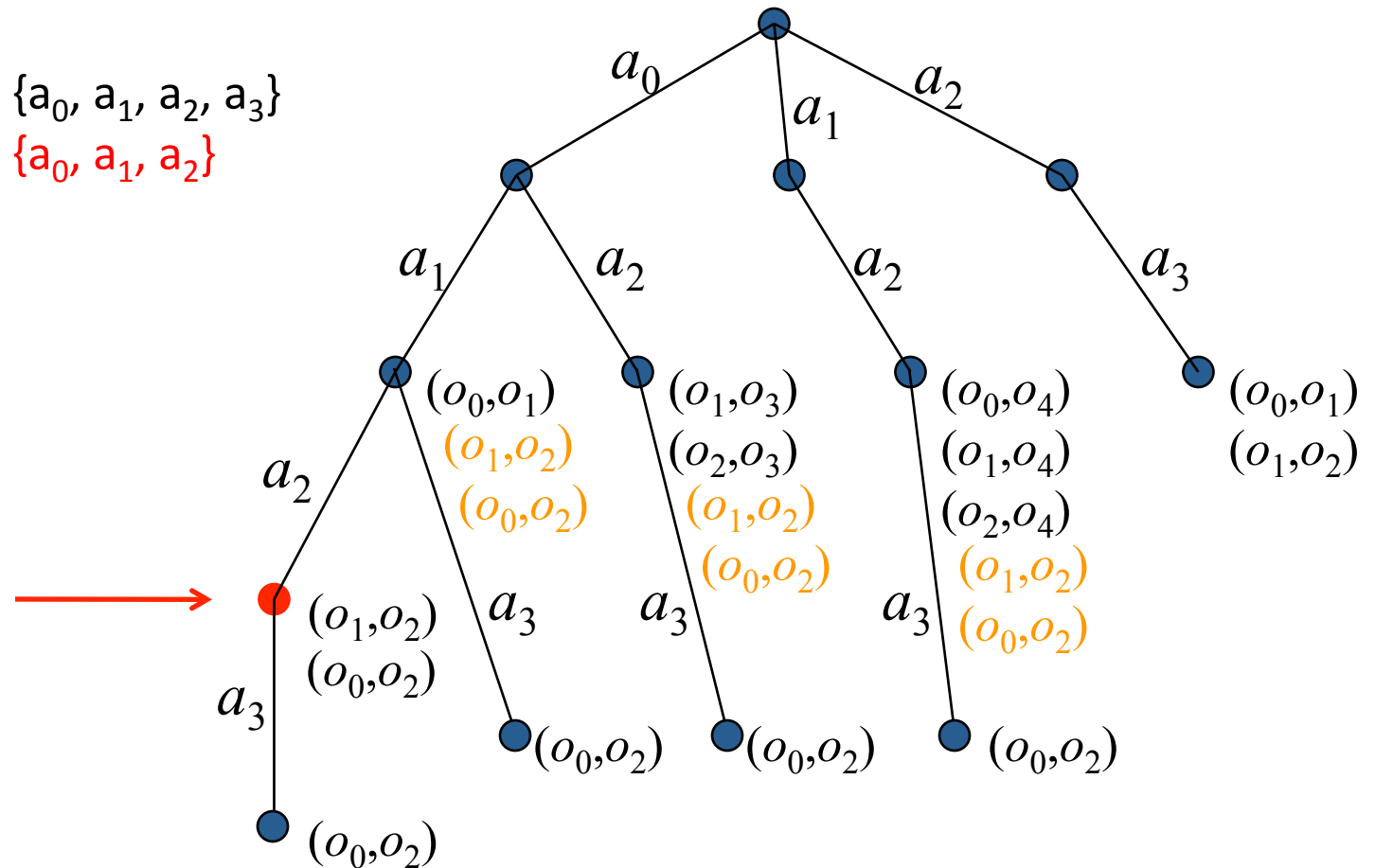
to

$$\begin{aligned} &\{a_0, a_1\}, \\ &\{a_0, a_2\}, \\ &\{a_1, a_2\} \end{aligned}$$


Coherent Cluster

4. Use **post-order traverse** the tree to find all maximum coherent clusters

- **Report:**

$$\{O_0, O_2\} \times \{a_0, a_1, a_2, a_3\}$$
$$\{O_1, O_2\} \times \{a_0, a_1, a_2\}$$


Coherent Cluster

4. Use **post-order traverse** the tree to find all maximum coherent clusters

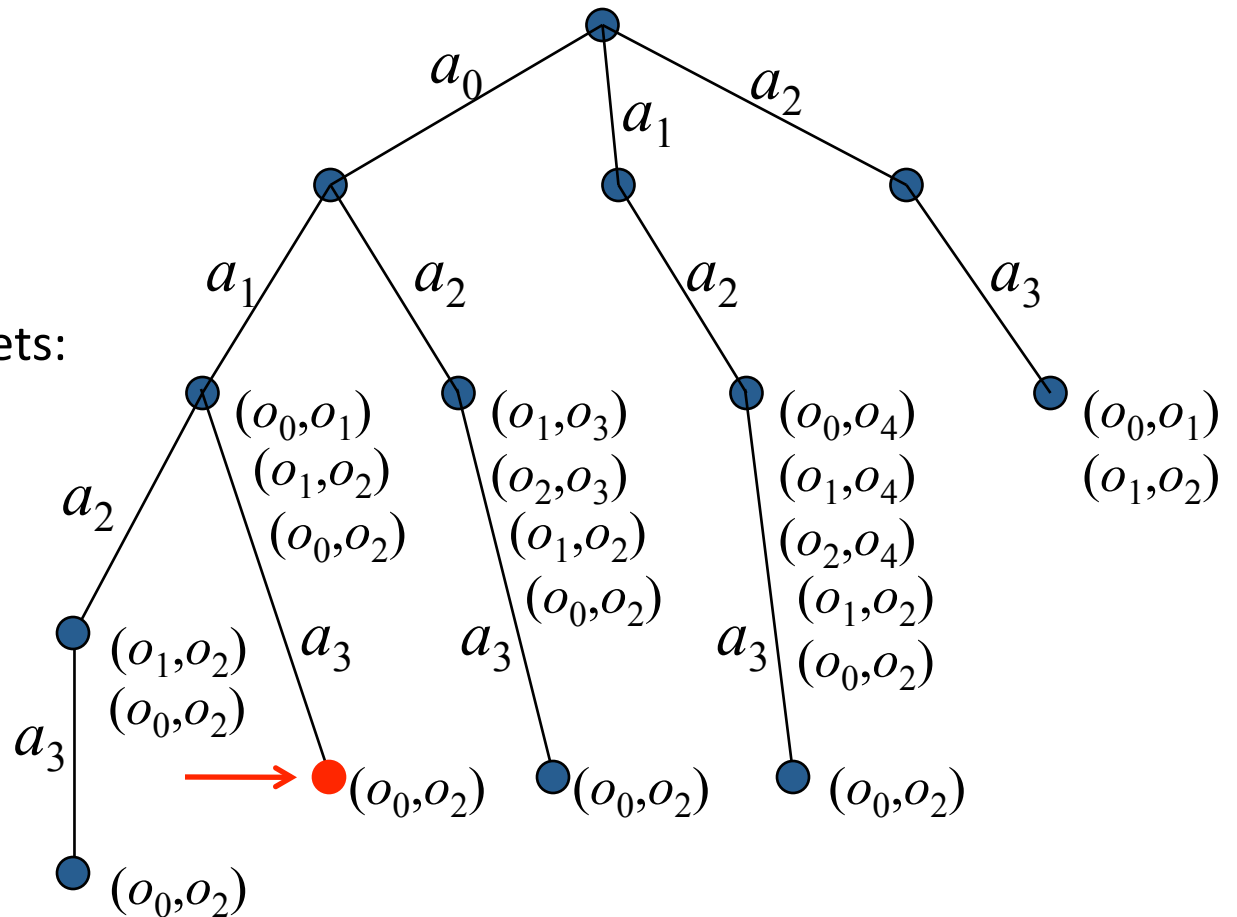
- Coherent object sets:

$$(O_0, O_2)$$

- Coherent attribute set:

$$\{a_0, a_1, a_3\}$$

- (K-1) Coherent attribute sets:

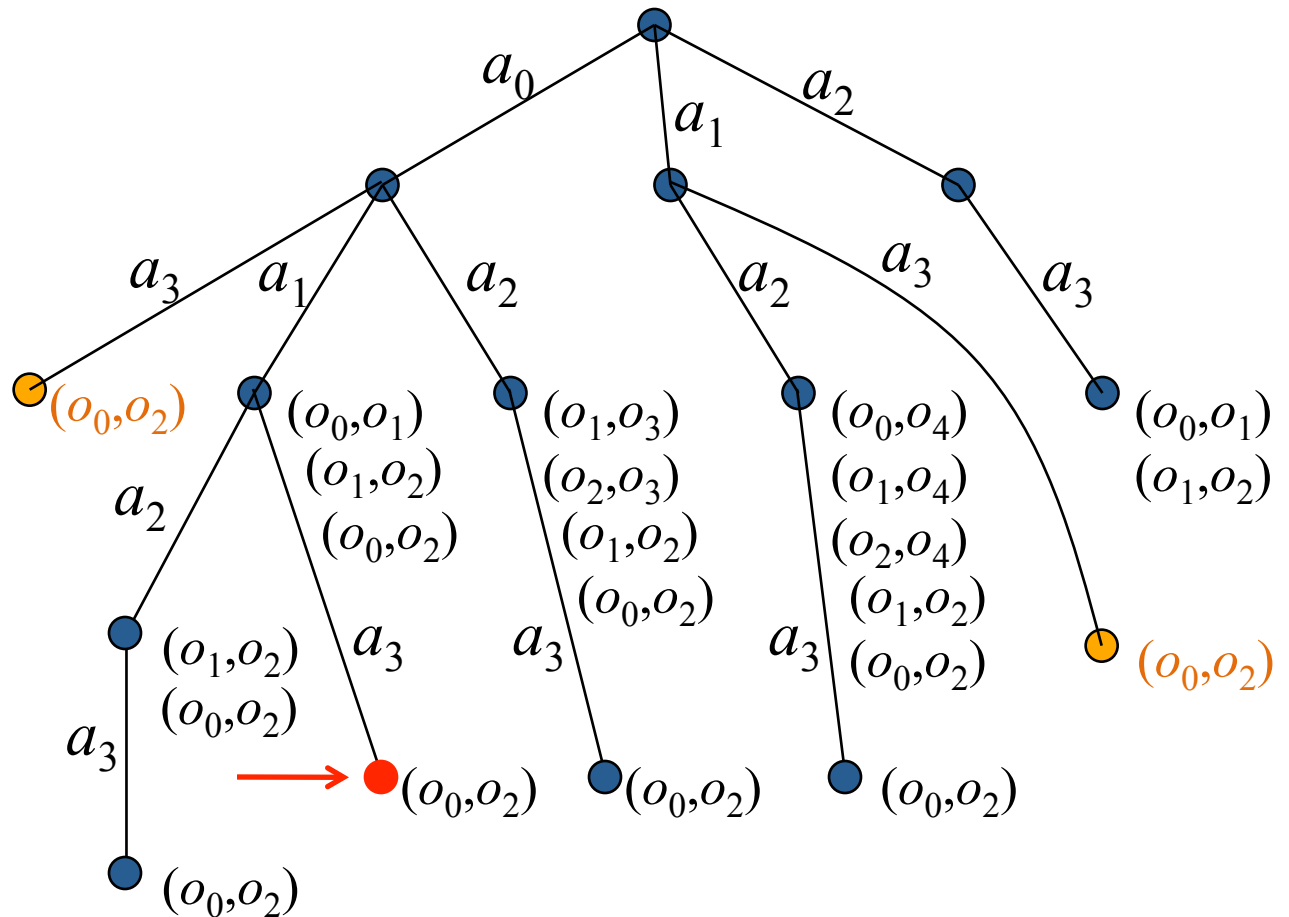
 $\{a_0, a_1\},$
$$\{a_0, a_3\}$$
$$\{a_1, a_3\}$$


Coherent Cluster

4. Use **post-order traverse** the tree to find all maximum coherent clusters

- **Copying**

to (O_0, O_2)

$$\begin{aligned} &\{a_0, a_1\}, \\ &\{a_0, a_3\}, \\ &\{a_1, a_3\} \end{aligned}$$


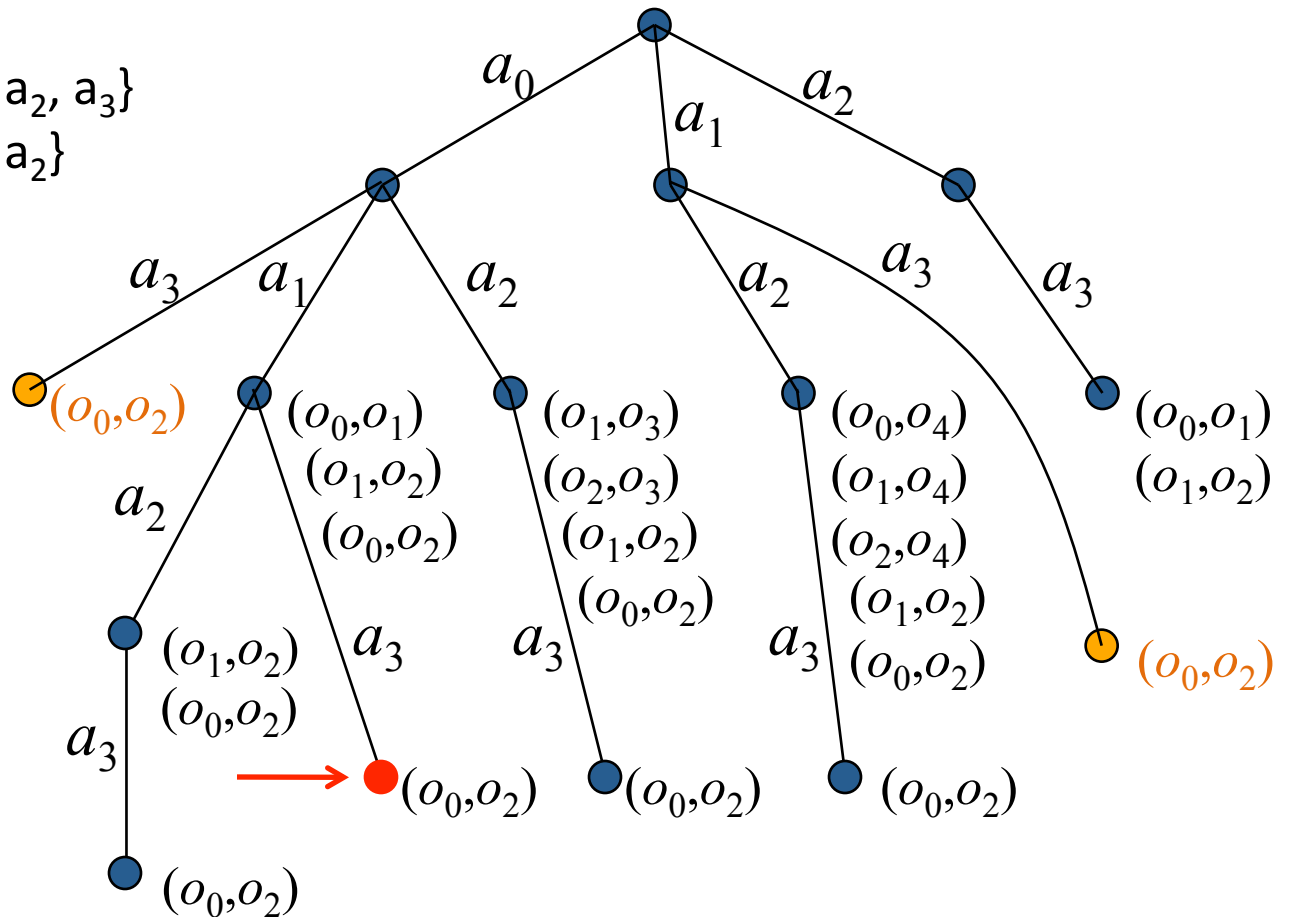
Coherent Cluster

4. Use **post-order traverse** the tree to find all maximum coherent clusters

- **Report:**

$$\{O_0, O_2\} \times \{a_0, a_1, a_2, a_3\}$$
$$\{O_1, O_2\} \times \{a_0, a_1, a_2\}$$
$$\{O_0, O_2\} \times \{a_0, a_1, a_3\}$$

is a subset of

$$\{O_0, O_2\} \times \{a_0, a_1, a_2, a_3\}$$


$$\{o_1, o_2\} \times \{a_0, a_1, a_2\}$$

$$\{o_0, o_1, o_2\} \times \{a_0, a_1\}$$

$$\{o_1, o_2, o_3\} \times \{a_0, a_2\}$$

$$\{o_0, o_2, o_4\} \times \{a_1, a_2\}$$

$$\{o_1, o_2, o_4\} \times \{a_1, a_2\}$$

$$\{o_0, o_1, o_2\} \times \{a_2, a_3\}$$

