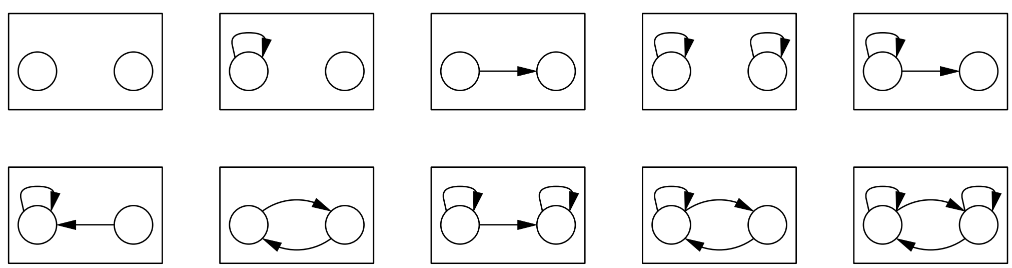
HOW TO DISTINGUISH UNLABELED DIGRAPHS

(n, m)-digraph. A digraph with $n \geq 1$ nodes and $m \geq 0$ links.

Need to Distinguish Two Unlabeled (n, m)-Digraphs.

- We use labeled digraphs to represent relations and unlabeled digraphs to represent structures of relations.
- To count #(structures of relations with some given properties), we need to know when two unlabeled (n, m)-digraphs are to be considered equal or, equivalently, different.
- We can easily distinguish the unlabeled digraphs shown below for structures of relations with the property that it is on a set of size $n = 2$ items by looking at them visually. But for large n and m, we need a more systematic method because the digraphys can be drawn in many different ways.

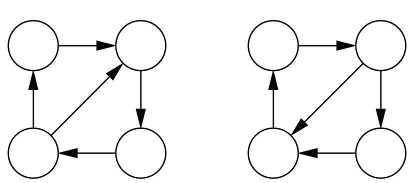


Equality of Two Unlabeled (n, m)-Digraphs.

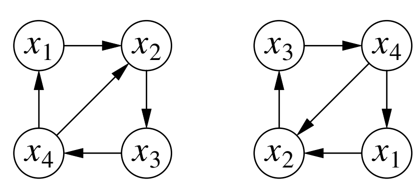
- They are equal if the nodes in each digraph can be labeled as x_1, x_2, $\cdots$, and x_n such that the links in the two digraphs match. (Fix a labeling for one digraph and try all $n!$ labelings of the other digraph to find one, if any, that matches the links.)

Example.

- The two unlabeled (4,5)-digraphs on the left seem to be different because the diagonal links in them are in opposite directions. However, their labeled versions shown on the right indicate that they are actually equal.



(i) A pair of (4,5)-unlabeled digraphs.



(ii) A labelling of digraphs in (i) that shows they are equal.

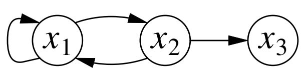
HOW TO DISTINGUISH TWO UNLABELED DIGRAPHS (Contd.)

Some Relevant Terminology

We define them first for labeled digraphs; some of them can be used for unlabeled digraphs as well.

(a) Outdegree(x_i) = outdegree of node x_i is #(links from x_i to other nodes or to itself).

(b) Outdegree-sequence(G) = the list of outdegree(x_i) of the nodes x_i in G, arranged in non-increasing order. It is independent of node-labels and can be used for unlabeled digraphs.



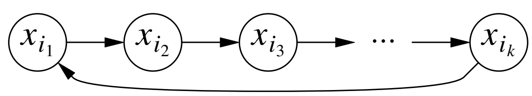
A labeled digraph G.

- Outdegree(x_3) = 0; outdegree-sequence = $\langle 2, 2, 0\rangle$.
- Indegree(x_3) = 1; indegree-sequence = $\langle 2, 1, 1\rangle$.
- 4 = #(links) = 2 + 2 + 0 = sum of outdegrees = 2 + 1 + 1 = sum of indegrees.
- The only cycle is $C = \langle x_1, x_2, x_1\rangle$ and length(C) = 2.

(c) Indegree(x_i) and indegree-sequence(G) are defined in a similar way. Note that $\sum$ outdegree(x_i) = $\sum$ indegree(x_i) = #(links in G).

(d) A path of length $k \geq 1$ is a list $\langle x_{i_1}, x_{i_2}, \cdots, x_{i_k}, x_{i_{k+1}}\rangle$ of $k+1$ distinct nodes, where each $(x_{i_j}, x_{i_{j+1}})$ is a link in the digraph; length = #(links in path).

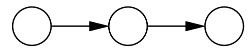
(e) A cycle of length $k \geq 2$ is a list $\langle x_{i_1}, x_{i_2}, \cdots, x_{i_k}, x_{i_1}\rangle$ of $k+1$ nodes, which are distinct except that the last node $x_{i_{k+1}}$ equals the first node x_{i_1}; length = #(links in cycle).



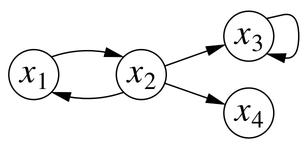
Note that visiting the same nodes $\{x_{i_1}, x_{i_2}, \cdots, x_{i_k}\}$ in the same cyclic-order but starting at a different node x_{i_j} is not considered a different cycle.

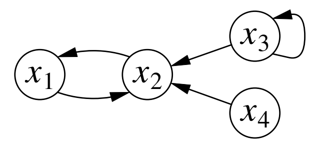
Practice Questions.

1. What different roles are played by labeled and unlabeled digraphs in the study of relations? Explain why an unlabeled digraph like the one below has no matrix representation.

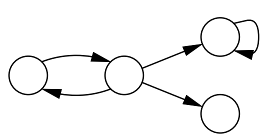


2. Why distinguishing two unlabeled (n, m)-digraphs is difficult when n and m are large? What is the maximum m = #(links) for a relation on R on a set of size n = #(nodes in the digraph)?
3. Show paths of length k and #(paths of length k) for $1 \leq k < 4$, for the digraphs below.



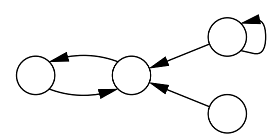


4. What is the maximum length of a path (cycle) in an (n, m)-digraph?
5. Is #(paths of length k) independent of node-labels for each $k \geq 1$? How about #(cycles of length k)?
6. Is #(cycles of length k) related to #(paths of length k) or to #(paths of length $k - 1$)?
7. How is #(paths of length k) related to #(paths of length $k + 1$)? Does a similar relation hold for #(cycles of length k) and #(cycles of length $k + 1$)?
8. Let the reverse-digraph $r(G)$ of G be the digraph obtained by reversing the direction of each link in G; see below. Clearly, reversing $r(G)$ gives back G. If G is the structure of a symmetric relation, then $r(G) = G$. We can also have $G = r(G)$ for non-symmetric relations; for example, G is a path, a cycle, a node disjoint union of two of more G_i's such that each $r(G_i) = G_j$ for some G_j in the union (j may be same as i). If G is the union of the two digraphs below, with a total of $2\times4 = 8$ nodes and $2\times5 = 10$ links, then $G = r(G)$.



An (unlabeled) digraph G

Its reverse-digraph $r(G)$



   If G is an unlabeled digraph, which of the following are the same for G and $r(G)$? In the other cases, give an example G to show that G and $r(G)$ differ in that respect.

   (a) Outdegree-sequence.

   (b) Indegree-sequence.

   (c) #(paths of length k), $k \geq 1$.

   (d) #(cycles of length k), $k \geq 2$.

9. If we know outdegree-sequence(G), what can you determine about $r(G)$?

HOW TO DISTINGUISH TWO UNLABELED DIGRAPHS (Contd.)

Distinguishing Two Unlabeled (n, m)-Digraphs.

The equality-test of two unlabeled digraphs defined earlier is not practical ($n!$ is too large even for a small n). A simpler way, though not a full-proof, is to test that the two digraphs have:

(E.1) The same number of links that connect a node to itself (they are called loops).

(E.2) The same maximum and minimum outdegrees and, likewise, for indegrees.

(E.3) The same outdegree-sequence and the same indegree-sequence. (This condition implies condition (E.2).)

(E.4) The same #(cycles of length k) for each $k \geq 2$ and, likewise, for paths.

Loops are not part of any cycle or path. Also, because #(paths of length 1) = m −#(loops), condition (E.1) implies #(paths of length 1) will be the same for two (n, m)-digraphs.

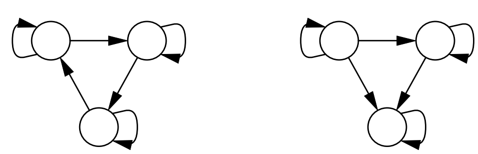
Limitation of Equality-Tests (E.1)-(E.4).

- Failing any of conditions (E.1)-(E.4) means the digraphs are not equal and represent different structures.
- The opposite is not true, passing all of (E.1)-(E.2) does not mean the unlabeled digraphs are equal.

In this sense, tests (E.1)-(E.4) are good only for detecting "not equal"

Example.

The (3, 6)-unlabeled digraphs below are not equal; they pass test (E.1) but fail tests (E.2)-(E.4).



Lack of an Efficient Equality-Test for Unlabeled Digraphs.

- One can have many other tests like (E.1)-(E.4) that are good for detecting that two (n, m)-digraphs are "not equal".
- However, there is no efficient method known at present to test the equality of two (n, m)-digraphs.

Practice Questions.

1. For the two unlabeled digraphs below, show that each of the equality-test (E.1)-(E.4) holds (and thus they cannot be distinguished based on those equality-tests).

2. Explain why the digraphs in Problem 1 are not considered equal, i.e., why "their nodes cannot be labeled as x_1, x_2, and x_3 such that $\cdots$".
3. Show the reverse-digraphs of the digraphs in Problem 1 and show that they cannot be distinguished by equality-tests (E.1)-(E.4).

   Also, argue that if two unlabeled (n, m)-digraphs cannot be distinguished by equality-tests (E.1)-(E.4), then the same is true for their reverse-digraphs.
5. If two unlabeled digraphs are equal, argue that their reverse-digraphs are also equal (by giving a proper labeling scheme for the reverse-digraphs).
6. Consider the definitions below.
   - Let the (total) degree of a node x_i be degree(x_i) = outdegree(x_i) + indegree(x_i), and the degree-sequence of an (unlabeled) digraph G be the list of degree(x_i)s of the nodes x_i in G arranged in the usual non-increasing order.

     The degree-sequences of the digraphs in Problem 1 are: $\langle 3, 2, 1, 0\rangle$ and $\langle 2, 2, 1, 1\rangle$.
   - Let the degreePair of a node x_i be degreePair(x_i) = (outdegree(x_i), indegree(x_i)), and the degreePair-sequence of an (unlabeled) digraph G be the list of degreePair(x_i) of the nodes x_i in G arranged first in the usual non-increasing order of outdegrees and then the pairs with the same outdegree arranged in non-increasing order of indegrees.

     The degreePair-sequences for the digraphs in Problem 1 are: $\langle(2,1), (1,0), (0,1), (0,1)\rangle$ and $\langle(2,0), (1,1), (0,1), (0,1)\rangle$.

   State two equality-tests (E.5) and (E.6) based on the above definitions. (The digraphs in Problem 1 can be distinguished based on each of (E.5)-(E.6).)
7. Give two unlabeled (n, m)-digraphs that passes equality-test (E.5) but not (E.6). (This shows (E.6) is stronger than (E.5), i.e., passing test (E.6) implies passing (E.5).)
8. Give two (n, m)-unlabeled digraphs that passes equality-tests (E.5)-(E.6) but not (E.4).
9. How can we compute the outdegree-sequence and indegree-sequence from the matrix-representation of a digraph?