CS40 - Summer 2013 : Notes

July 19
Victor Amelkin

Set Theory

① Basic definitions

def A set is an unordered collection of unique objects.

Ex 1: $S = \{1, \text{"Alice"}, ☺\}$ - a set.

Yet S in the example above is a legal set (according to the given definition), it is not typical in that its elements are of different nature. We often deal with sets of similar objects (objects of the same "type"):

$$W = \{1, 3, 7.5, 1/3\}.$$

Ex 2: $S = \{1, 1, 2\}$ - not a set (repetitions are not allowed).

Ex 3: $A = \{1, 2\}$, $B = \{2, 1\}$ - sets A and B are the same (order does not matter).

def If element x belongs to set S, we say that "x is in S" and write $x \in S$.
If element x does not belong to set S, we write $x \notin S$ (though, $\neg x \in S$ is also OK :)).
Given any object x, it should either belong to a set S or not to; if this is not the case (i.e., both options are not available, or both can hold simultaneously, then S is not a set).

Note: $x \in S$ and $x \notin S$ are the only two options, and only one of them can be true for given x and S. But, it is not always possible to tell whether $x \in S$. Sets S for which we can always tell whether $x \in S$ (for any x) are called decidable. All sets we are going to work with are such.

Sets can contain other sets:

$A = \{\{1, 2\}, \{3, 4\}\}$ - a set containing two elements - $\{1, 2\}$ and $\{3, 4\}$ -, both of which also happen to be sets.

$\{1, 2\} \in A$, $\{3, 4\} \in A$

$1 \notin A$

def: Empty set, denoted with $\emptyset$, is a set containing no elements.

② Ways to define a set

1) Listing all elements:

$S = \{A, B, C, D\}$

- only works for finite sets;
- and only if the "size" of a set is reasonably small.

2) Describing sets informally:

$S = \{$all even natural numbers$\}$ (that is, $S = \{2, 4, 6, 8, \dots\}$)

- works for "large" sets, but not very formal; we cannot do formal reasoning or algebra on such-defined sets.

! 3) Describing a set with a predicate ("set builder"):

$A = \{ i \underbrace{\,|\,}_{\text{"such that"}} \exists x \in \mathbb{N} : i = x^2 \}$ (i.e., $A = \{1, 4, 9, 16, \dots\}$)

$B = \{ 2^x \mid x \text{ is prime} \}$ (i.e., $B = \{2^2 = 4, 2^3 = 8, 2^5 = 32, 2^7 = 128, \dots\}$)

$C = \{ 1 \mid x \text{ is prime} \}$ (i.e., $C = \{1\}$)

③ Naïve set theory and Russel's paradox

The naïve set theory is based on the definition of a set we gave. This definition is very simple, and it allows us to define "sets" that look like sets, but does not behave as such. One of such "sets" was proposed by Bertrand Russel:

$S = \{ x \mid x \notin x \}$, that is S contains elements which are not members of themselves (it is assumed that S's elements are sets).

Suppose $S \in S$. Then, by definition of S, $S \notin S$ — a contradiction.

Suppose $S \notin S$. Then, it is not true that $S \notin S$, i.e., $S \in S$ — again, a contradiction.

Thus, S is not a set, since for a given element (namely, S), none of two options $\in$ and $\notin$ is possible.

Such "tricky" sets are handled by axiomatic set theory; naïve set theory works for "normal" sets very well.

④ Cardinality of a set

def For a *finite* set A, its *cardinality* (or size) is the number of elements it has.

Ex: $A = \phi$, $\overbrace{|A|} = 0$

$B = \{a, b, c\}$, $|B| = 3$

$C = \{\underbrace{\phi}, \underbrace{\{1,2\}}, \underbrace{\{\text{"Alice"}, \text{"Bob"}\}}\}$, $|C| = 1 + 1 + 1 = 3$

$D_n = \{s \mid s \text{ is a binary string of length } n\}$, $|D_n| = 2^n$

For infinite sets, the given definition of cardinality does not make much sense:

$\mathbb{N} = \{1, 2, 3, \dots\}$ $\quad |\mathbb{N}| = \infty$?

$\mathbb{R} = \{x \mid x \text{ is real}\}$ $\quad |\mathbb{R}| = \infty$?

$$|\mathbb{N}| \stackrel{?}{=} |\mathbb{R}|$$

For infinite sets, there are "cardinality classes":

1) *Countable sets* - infinite sets whose elements can be enumerated:

$\mathbb{N}$ - set of natural numbers $\{1, 2, 3, 4, \dots\}$

$\mathbb{Q}$ - set of rational numbers

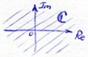
$$\left\{ \begin{array}{ccccc} \pm\frac{1}{1} & \pm\frac{1}{2} & \pm\frac{1}{3} & \pm\frac{1}{4} & \dots \\ \pm\frac{2}{1} & \pm\frac{2}{2} & \pm\frac{2}{3} & \pm\frac{2}{4} & \dots \\ \pm\frac{3}{1} & \pm\frac{3}{2} & \pm\frac{3}{3} & \pm\frac{3}{4} & \dots \\ \vdots & & & & \end{array} \right\}$$

$\mathbb{Z}$ - set of integers $\{0, \pm 1, \pm 2, \pm 3, \dots\}$

$\mathbb{Z}^+$ - set of non-negative integers $\{0, 1, 2, 3, \dots\}$

2) Sets with cardinality of the *continuum*:

$\mathbb{R}$ - set of real numbers 

$\mathbb{C}$ - set of complex numbers 

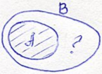
3) Sets with cardinality of *hypercontinuum*:

$2^{\mathbb{R}}$ - powerset of $\mathbb{R}$ (set of all subsets of $\mathbb{R}$)

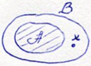
> "Continuum Hypothesis":
> [Georg Cantor, 1878]
>
> There is no set with cardinality larger than that of $\mathbb{N}$ and smaller than that of $\mathbb{R}$.
>
> - Cannot be proven or disproven in naïve or axiomatic set theories.

⑤ Equality and Containment of Sets

def $A \subseteq B$ ("A is contained in B") if $\forall x : [(x \in A) \rightarrow (x \in B)]$



def $A \subset B$ ("A is properly contained in B") if $(A \subseteq B) \wedge \exists x : (x \in B \wedge x \notin A)$



def $A = B$ ("Sets A and B are equal") if $A \subseteq B \wedge B \subseteq A$

(Negative versions $\not\subseteq$, $\neq$ are naturally defined.)

Note: If $A \subset B$ then $A \subseteq B$. If $A \subseteq B$, then not necessarily $A \subset B$.

Ex: $A = \{x \mid x \text{ is even natural number}\}$

$B = \{2^x \mid x \text{ is prime}\}$

$B \subseteq A$ (since 2^x is always an even natural number for prime x)

$B \subset A$ (since $2^4 = 16 \in A$, but $\notin B$ (since 4 is not prime))

def: If $A \subseteq B$, then A is a subset of B.

If $A \subset B$, then A is a proper subset of B

Note: Do not mix up $\in$ and $\subset$ (or $\subseteq$). $\in$ - for membership in a set

$\subset$ - for containment of sets

$\subseteq$ "

$S = \{\emptyset, \{1,2\}\}$

$\{1,2\} \in S$

$\{1,2\} \not\subset S$ $\quad \{\{1,2\}\} \subset S$

$\emptyset \in S$, $\emptyset \subset S$

def The powerset of set S is the set of all subsets of S. (Notation: $P(S)$)

Ex: $X = 1$ $\quad P(X) = ?$ (X is not a set, so $P(X)$ is undefined)

$X = \emptyset$ $\quad P(X) = \{\emptyset\}$

$X = \{\emptyset\}$ $\quad P(X) = \{\emptyset, \{\emptyset\}\}$

$X = \{\emptyset, \emptyset\}$ $\quad P(X) = ?$ (again, X is not a set)

$X = \{\emptyset, \{\emptyset\}\}$ $\quad P(X) = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}$

$X = \{1,2,3\}$ $\quad P(X) = \{\emptyset,$
$\{1\}, \{2\}, \{3\},$
$\{1,2\}, \{1,3\}, \{2,3\},$
$\{1,2,3\}$
$\}$

Hypothesis: $|P(X)| = 2^{|X|}$

(will be proven later)

⑥ Representing subsets; enumerating subsets; $|P(X)|$

Given a finite set S, we would like to enumerate (or print out) all its subsets. To do so, we need a convenient representation for subsets.

$S = \{a_1, a_2, a_3, \dots, a_{n-1}, a_n\}$

1	0	1	...	1	0

– binary string of length n

Each subset will be represented with a binary string of length n; if bit i in a string is 0, then a_i does not belong to the corresponding subset, and it does if the i'th bit is 1. For example,

$A = \{a_2, a_4\} \subset S \sim$

1	2	3	4	5	6		n
0	1	0	1	0	0	...	0

$B = \emptyset \subset S \sim$

1	2	3		n
0	0	0	...	0

$S \subseteq S \sim$

1	2	3	4	5		n-1	n
1	1	1	1	1	...	1	1

Thus, the number of subsets of set S is exactly the number of binary strings of length $|S|$. The same result can be obtained as follows:

$|P(S)| = (\text{let us denote } |S| \text{ as } n) = \underbrace{C(n,0)}_{\#\text{ of subsets of size } 0} + \underbrace{C(n,1)}_{\#\text{ of subsets of size } 1} + \underbrace{C(n,2)}_{\#\text{ of subsets of size } 2} + \dots + \underbrace{C(n,n)}_{\#\text{ of subsets of size } n} =$

$= \left(\text{use binomial theorem } (x+y)^n = \sum_{k=0}^{n} C_n^k x^k y^{n-k} \text{ with } x=y=1\right) = (1+1)^n = 2^n.$

Consequently, the algorithm for enumerating all subsets of a finite set may look as follows:

```
enum-subsets (set S) {
    // assert (S is finite)
    for i = 1 to 2^|S| {
        binstr = to_binary(i)
        subset = {}
        for j = 1 to |S| {
            if(binstr[j] == 1) {
                subset += S[j]
            }
        }
        print (subset)
    }
}
```

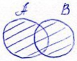
What if we need to enumerate only subsets of a fixed size k and we do not want to go through all the $2^{|S|}$ options? In other words, is there a way to efficiently enumerate all binary strings of length n with exactly k bits set to 1?

One solution is to use "snoob" bit hack:

http://hackersdelight.org/hdcodetxt/snoob.c.txt

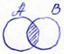
⑦ Operations on sets

Union $A \cup B$



$\{1,2,3\} \cup \{2,3,4,5\} \cup \emptyset = \{1,2,3,4,5\}$

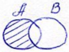
Intersection $A \cap B$



$\{1,2,3\} \cap \{2,4,9\} = \{2\}$

$\{a,b\} \cap \{c,d\} = \emptyset \quad (\{a,b\} \not\cap \{c,d\})$

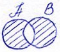
Difference $A \setminus B$



$\{1,2,3,4,5\} \setminus \{1,3,5\} = \{2,4\}$

$\{1,2,3\} \setminus \{9,10\} = \{1,2,3\}$

Symmetric difference ("XOR") $A \Delta B$

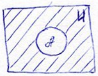


$\{1,2,3\} \Delta \{3,4,5\} = \{1,2,4,5\}$

$A \Delta B = (A \setminus B) \cup (B \setminus A)$

$\quad\quad = (A \cup B) \setminus (A \cap B)$

Complementation $\overline{A}$

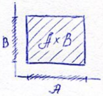


U - universe - should be explicitly defined in order to use complementation.

$U = \mathbb{N}$, $A = \{x \mid x \text{ is natural and even}\}$

$\overline{A} = \{x \mid x \text{ is natural and odd}\}$

Cartesian product $A \times B$
(cross-product)

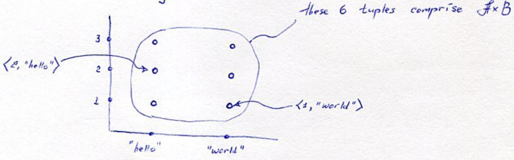


(Generalizes to $A_1 \times A_2 \times \ldots \times A_n$)

$A = \{1,2,3\}$, $B = \{\text{"hello"}, \text{"world"}\}$

$A \times B = \{$
$\langle 1, \text{"hello"} \rangle,$
$\langle 1, \text{"world"} \rangle,$
$\langle 2, \text{"hello"} \rangle,$
$\langle 2, \text{"world"} \rangle,$
$\langle 3, \text{"hello"} \rangle,$
$\langle 3, \text{"world"} \rangle$
$\}$

($\langle x,y \rangle$ - (ordered) tuple
$\langle 1,2 \rangle \neq \langle 2,1 \rangle$)



⑧ Laws of set theory

$A \cap \emptyset = \emptyset$, $A \cup \emptyset = A$ } empty set laws

$A \cap A = A$, $A \cup A = A$ } idempotency laws

$A \cup B = B \cup A$, $A \cap B = B \cap A$ } commutative laws

$A \cap (B \cap C) = (A \cap B) \cap C$, $A \cup (B \cup C) = (A \cup B) \cup C$ } associative laws

$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ } distributive laws

$A \cap (A \cup B) = A$, $A \cup (A \cap B) = A$ } absorption laws

De Morgan's laws:

$$A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$$

$$A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$$

or, using complementation

$$\overline{B \cap C} = \overline{B} \cup \overline{C}$$

$$\overline{B \cup C} = \overline{B} \cap \overline{C}$$

laws related to the operation of complementation:

$$\overline{\overline{A}} = A$$

$$\overline{\emptyset} = U$$

$$A \cap \overline{A} = \emptyset$$

$$A \cup \overline{A} = U$$

Note: To prove equalities for sets, either use the laws of set theory to derive a chain of equivalences or use truth tables:

$x \in A$	$x \in B$	$x \in A \cup B$	$x \in \overline{\overline{A} \cap \overline{B}}$
0	0	0	0
0	1	1	1
1	0	1	1
1	1	1	1

- in reality, we would als[o] have individu[al] columns for sub-expression[s] like $\overline{A}$.

⑨ Cardinality of "complex" sets:

$|A \cup B| = |A| + |B| - |A \cap B|$ (inclusion-exclusion for 2 sets)

$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$ (— " — for 3 sets)

$|A| = |U| - |\overline{A}|$ – works only for finite sets

$|A \times B| = |A| \cdot |B|$

Exercise: Prove that $|\overline{A} \cap \overline{B}| = |U| - |A| - |B| + |A \cap B|$, ($A, B$ – finite sets, U – universe).

Proof: $|\overline{A} \cap \overline{B}| =$ (De Morgan's law) $= |\overline{A \cup B}| =$ (cardinality of complement of a finite set) $= |U| - |A \cup B| =$
$=$ (inclusion-exclusion) $= |U| - |A| - |B| + |A \cap B|$. □

(10) Relations

$A_1, A_2, \ldots, A_n$ - sets

def $\langle a_1, a_2, \ldots, a_n\rangle$ - n-tuple (or just tuple) on $A_1, \ldots, A_n$ - ordered sequence of elements $a_i \in A_i$, $i=1,\ldots,n$.

Ex: $A_1 = \{1, 2\}$, $A_2 = \{\text{Alice}, \text{Bob}\}$ $(n=2)$

$\langle 2, \text{Alice}\rangle$ - 2-tuple ("pair") on A_1, A_2.

$\langle 2, \text{Alice}\rangle \neq \langle \text{Alice}, 2\rangle$ - order matters.

$A_1 \times A_2 \times \ldots \times A_n = \{\langle a_1, a_2, \ldots, a_n\rangle \mid a_i \in A_i,\ i=1,\ldots,n\}$ - cross-product (Cartesian product) of $A_1, A_2, \ldots, A_n$ - the set of all the possible tuples on $A_1, \ldots, A_n$.

def A relation R on $A_1, A_2, \ldots, A_n$ is a (not necessarily proper) subset of $A_1 \times A_2 \times \ldots \times A_n$.

Ex: $A_1 = \{1\}$, $A_2 = \{\text{Alice}, \text{Bob}\}$, $A_3 = \{\bigcirc, \triangle, \square\}$ $(n=3)$

$R_0 = \emptyset$ - a legal relation on A_1, A_2, A_3 (in fact, R_0 is a legal relation on any sets as long as R_0's arity (defined below) has not been pre-defined)

$R_1 = \{\langle 1, \text{Alice}\rangle\}$ - a relation on A_1, A_2.

$R_2 = \{\langle 1, \bigcirc, \text{Alice}\rangle, \langle 1, \triangle, \text{Bob}\rangle\}$ - a relation on A_1, A_3, A_2.

$R_3 = \{\langle 1, 1, 1\rangle\}$ - a relation on A_1, A_1, A_1.

$R_4 = \{\langle 1, \bigcirc\rangle, \langle \bigcirc, 1\rangle\}$ - not a relation on A_1, A_3; not a relation on A_3, A_1

def: Arity of relation $R \subseteq A_1 \times A_2 \times \ldots \times A_n$ is n. When $n=2$, relation is "binary", and its tuples are called "pairs". When $n=1$, a relation is "unary", or we can just treat it as a subset ($R \subseteq A_1$). Arity can be 0; a relation of such arity can either be empty ($\emptyset$) or or contain an empty tuple $\langle\rangle$.

def: Cardinality (or size) of a relation is defined as the set cardinality (relations are sets).

Ex: $R \subseteq \text{People} \times \mathbb{N} \times \mathbb{N} \times \mathbb{N}$

$R = \{\langle p, d, m, y\rangle \mid \text{person } p\text{'s birth date is } m\text{-}d\text{-}y\}$

$\langle \text{me}, 29, 8, 1985\rangle \in R$

$\text{arity}(R) = 4$

$|R| = |\text{People}|$ (cardinality)

Membership in a relation:

$\langle a_1, a_2, \dots, a_n \rangle \overset{\notin}{\in} R \; (\subseteq A_1 \times A_2 \times \dots \times A_n)$ – most general notation.

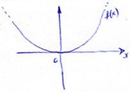
$R(a_1, a_2, \dots, a_n)$ – "predicate notation" (how to express $\notin$ using this notation?)

$a_1 R a_2 \; (\sim \langle a_1, a_2 \rangle \in R)$
$a_1 \not R a_2 \; (\sim \langle a_1, a_2 \rangle \notin R)$
> special notation for binary relations.

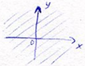
Functions as relations:

$f: \mathbb{R} \to \mathbb{R}$

$f(x) = x^2$



$\mathbb{R} \times \mathbb{R} = \{\langle x, y \rangle \mid x \in \mathbb{R}, y \in \mathbb{R}\}$ –



(2d-plane)

$f \subseteq \mathbb{R} \times \mathbb{R}$

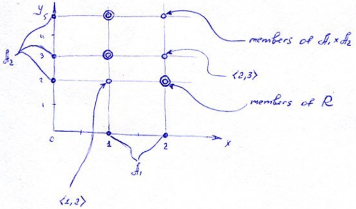
$f = \{\langle x, y \rangle \mid x \in \mathbb{R}, y \in \mathbb{R}, y = x^2\}$

Representing relations:

1) Graph ("grid" – a better name when relations are defined on discrete sets):

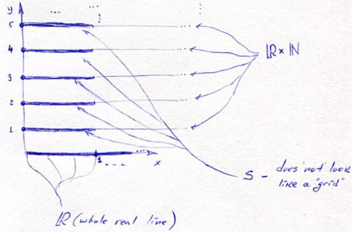
$R \subseteq \underset{A_1}{\{1, 2\}} \times \underset{A_2}{\{2, 3, 5\}}$

$R = \{\langle x, y \rangle \mid x \in A_1, y \in A_2, (x+y) \text{ – even}\}$



$S \subseteq \mathbb{R} \times \mathbb{N}$

$S = \{\langle x, y \rangle \mid x \in \mathbb{R}, y \in \mathbb{N}, 0 \le x \le 1, y \le 5\}$



2) Boolean matrix:

$A_1 \backslash A_2$	2	3	5
1	0	1	1
2	1	0	0

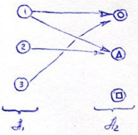
$\langle 1, 5 \rangle \in R$

$\langle 2, 3 \rangle \notin R$

3) Directed graph:

$R \subseteq \underset{A_1}{\{1, 2, 3\}} \times \underset{A_2}{\{\bigcirc, \triangle, \square\}}$

$R = \{\langle 1, \bigcirc \rangle, \langle 1, \triangle \rangle, \langle 2, \triangle \rangle, \langle 3, \bigcirc \rangle\}$



Combining relations:

Relations – sets ⇒ can do $\cup, \cap, \setminus, \Delta, \dots$

$$R_1 = \begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline 1 & 1 & 0 & 0 \\ 2 & 0 & 0 & 1 \\ 3 & 0 & 1 & 0 \end{array} \qquad R_2 = \begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline 1 & 1 & 1 & 1 \\ 2 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \end{array} \qquad R_1, R_2 \subseteq \underbrace{A \times A}_{\{1,2,3\}}$$

$R_1 \cup R_2 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ (~"disjunction")

$R_1 \cap R_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ (~"conjunction")

$R_1 \setminus R_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ (~"$\max\{x-y, 0\}$")

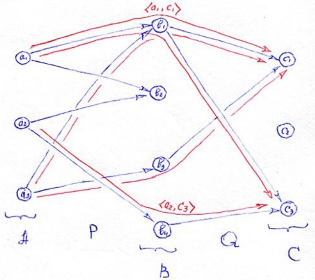
$R_1 \Delta R_2 = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ (~"XOR")

Composition of relations:

A, B, C – sets

$\left.\begin{array}{l} P \in A \times B \\ Q \in B \times C \end{array}\right\}$ relations

$$\underbrace{P \circ Q}_{\text{composition of } P \text{ and } Q} = \{ \langle a, c \rangle \mid \exists b : \underbrace{P(a,b)}_{\langle a,b \rangle \in P} \wedge \underbrace{Q(b,c)}_{\langle b,c \rangle \in Q} \}$$



$P = \{ \langle a_1, b_1 \rangle, \langle a_1, b_2 \rangle, \langle a_2, b_2 \rangle, \langle a_2, b_4 \rangle, \langle a_3, b_1 \rangle, \langle a_3, b_3 \rangle \}$

$Q = \{ \langle b_1, c_1 \rangle, \langle b_1, c_3 \rangle, \langle b_3, c_1 \rangle, \langle b_4, c_3 \rangle \}$

$P \circ Q = \{ \langle a_1, c_1 \rangle, \langle a_1, c_3 \rangle, \langle a_2, c_3 \rangle, \langle a_3, c_1 \rangle, \langle a_3, c_3 \rangle \}$

To compose $P \circ Q$, look for paths of length 2 from P to Q.

Self-study: – composition of binary relations
– boolean matrix multiplication over $\langle \overset{(\times)}{\wedge}, \overset{(+)}{\vee} \rangle$ – semiring.

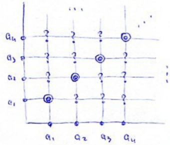
Properties of relations:

R - binary relation on A. $(R \subseteq A \times A)$

1) R - reflexive iff $\forall a \in A: \langle a,a\rangle \in R$

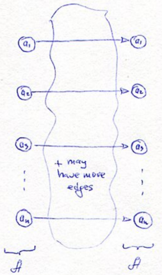
A \ A	a_1	a_2	a_3	...	a_n
a_1	1	?	?	...	?
a_2	?	1	?	...	?
a_3	?	?	1	...	?
⋮	⋮	⋮	⋮	⋱	?
a_n	?	?	?	...	1

Boolean matrix



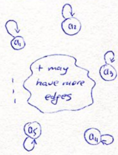
Grid

Directed graph:



OR

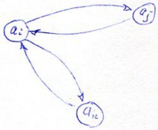
Compact notation (only one copy of A):



2) R - symmetric iff $\forall a,b \in A: \langle a,b\rangle \in R \to \langle b,a\rangle \in R$

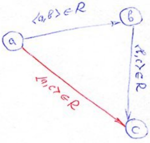
A \ A	a_1	a_2	a_3	...
a_1		x	y	⋮
a_2	x		z	⋮
a_3	y	z		⋮
⋮	...	...	...	⋱

Boolean matrix



Directed graph

3) R - transitive iff $\forall a,b,c \in A: (\langle a,b\rangle \in R \wedge \langle b,c\rangle \in R) \to (\langle a,c\rangle \in R)$



Question: symmetric + transitive $\overset{?}{\Rightarrow}$ reflexive? (Try

	a_1	a_2	a_3
a_1	1	0	0
a_2	0	0	0
a_3	0	0	1

)

Self-study: closures, equivalence relations, partitioning into equivalence classes.