Closures of Relations

Chapter 10
22c:19
Hantao Zhang

Relational closures

- Three types we will study
  - Reflexive
    - Easy
  - Symmetric
    - Easy
  - Transitive
    - Hard

Reflexive closure

- Consider a relation \( R \). We want to add edges to make the relation reflexive.
  
  \[ R' = R \cup \{(x, y)\} \]

- With matrices, we set the diagonal to all 1's

- By adding those edges, we have made a non-reflexive relation \( R \) into a reflexive relation

- This new relation \( R' \) is called the reflexive closure of \( R \)
Reflexive closure example

- Let \( R \) be a relation on the set \( \{0, 1, 2, 3\} \) containing the ordered pairs \((0,1), (1,1), (1,2), (2,0), (2,2), \) and \((3,0)\)
- What is the reflexive closure of \( R \)?
- We add all pairs of edges \((a,a)\) that do not already exist

Symmetric closure

- In order to find the symmetric closure of a relation \( R \), we add an edge from \( a \) to \( b \), where there is already an edge from \( b \) to \( a \)
- The symmetric closure of \( R \) is \( R \cup R^{-1} \)
  - If \( R = \{(a,b) \mid \ldots\} \)
  - Then \( R^{-1} = \{(b,a) \mid \ldots\} \)

Symmetric closure example

- Let \( R \) be a relation on the set \( \{0, 1, 2, 3\} \) containing the ordered pairs \((0,1), (1,1), (1,2), (2,0), (2,2), \) and \((3,0)\)
- What is the symmetric closure of \( R \)?
- We add all pairs of edges \((a,b)\) where \((b,a)\) exists
  - We make all "single" edges into anti-parallel pairs

We add edges:
- \((0,2), (0,3)\)
- \((1,0), (2,1)\)
Transitive closure

Given the flight relation between cities, what’s the relation “reachable by air” between cities?

The transitive closure would contain edges between all nodes reachable by a path of any length.

Transitive closure

- Informal definition: If there is a path from $a$ to $b$, then there should be an edge from $a$ to $b$ in the transitive closure.
- First take of a definition:
  - In order to find the transitive closure of a relation $R$, we add an edge from $a$ to $c$, when there are edges from $a$ to $b$ and $b$ to $c$.

$$R = \{(1,2), (2,3), (3,4)\}$$

1. $(1,2) \& (2,3) : (1,3)$
2. $(2,3) \& (3,4) : (2,4)$

Transitive closure

- Informal definition: If there is a path from $a$ to $b$, then there should be an edge from $a$ to $b$ in the transitive closure.
- Second take of a definition:
  - In order to find the transitive closure of a relation $R$, we add an edge from $a$ to $c$, when there are edges from $a$ to $b$ and $b$ to $c$.
  - Repeat this step until no new edges are added to the relation.
- We will study different algorithms for determining the transitive closure.

- **red** means added on the first repeat
- **teal** means added on the second repeat

1. $(1,2)$
2. $(2,3)$
3. $(1,3)$
4. $(2,4)$
6 degrees of separation

- The idea that everybody in the world is connected by six degrees of separation
  - Where 1 degree of separation means you know (or have met) somebody else
- Let $R$ be a relation on the set of all people in the world
  - $(a, b) \in R$ if person $a$ has met person $b$
- So six degrees of separation for any two people $a$ and $g$ means:
  - $(a, b), (b, c), (c, d), (d, e), (e, f), (f, g)$ are all in $R$
- Or, for any $a$ and $g$, $(a, g) \in R^n$

Connectivity relation

- $R$ contains edges between all the nodes reachable via 1 edge
- $R^2 = R \circ R$ contains edges between nodes that are reachable via 2 edges or less in $R$ (a path of length 2 or less)
- $R^n = R \circ R \circ \cdots \circ R$ contains edges between nodes that are reachable via $n$ edges or less in $R$ (a path of length $n$ or less)
- $R^*$ contains edges between nodes that are reachable via any number of edges (i.e. via any path) in $R$
  - Rephrased: $R^*$ contains all the edges between nodes $a$ and $b$ when is a path of length at least 1 between $a$ and $b$ in $R$
- $R^*$ is the transitive closure of $R$
  - The definition of a transitive closure is that there are edges between any nodes $(a,b)$ whenever there is a path from $a$ to $b$.

How long are the paths in a transitive closure?

- Let $R$ be a relation on set $A$, where $|A| = n$
  - Rephrased: consider a graph $G$ with $n$ nodes and some number of edges
- Lemma 1: If there is a path from $a$ to $b$ in $R$, then there is a path between $a$ and $b$ of length $< n$.
- Proof preparation:
  - Suppose there is a path from $a$ to $b$ in $R$
  - Let the length of that path be $m$
  - Let the path be edges $(x_0, x_1), (x_1, x_2), \ldots, (x_{m-1}, x_m)$
  - That’s $m+1$ nodes $x_0, x_1, x_2, \ldots, x_{m-1}, x_m$
  - If a node exists twice in our path, then it’s not a shortest path
    - As we made no progress in our path between the two occurrences of the repeated node
  - Thus, each node may exist at most once in the path, $m+1 \leq n$. 

10

11

12
How long are the paths in a transitive closure?

- Proof by contradiction:
  - Assume that the shortest path from a to b is at least \( n \), i.e., \( m \geq n \)
  - There are \( m+1 \) nodes in the path.
  - By the pigeonhole principle, there must be at least one node in the graph that has two occurrences in the path.
    - Not possible, as the path would not be the shortest path.
    - Thus, it cannot be the case that \( m \geq n \)

- If there exists a path from a to b, then there is a path from a to b of at most length \( n - 1 \).

Finding the transitive closure

- Let \( M_R \) be the zero-one matrix of the relation \( R \) on a set with \( n \) elements. Then the zero-one matrix of the transitive closure \( R^* \) is:
  \[
    M_{R^*} = M_R \lor M_R^{[2]} \lor M_R^{[3]} \lor \cdots \lor M_R^{[n-1]}
  \]

  Nodes reachable with one application of the relation
  Nodes reachable with two applications of the relation
  Nodes reachable with \( n-1 \) applications of the relation

Sample questions

- Find the zero-one matrix of the transitive closure of the relation \( R \) given by:

  \[
  M_R = \begin{bmatrix}
  1 & 0 & 1 \\
  0 & 1 & 0 \\
  1 & 1 & 0 \\
  \end{bmatrix}
  \]
  \[
  M_{R^*} = M_R \lor M_R^{[2]} = \begin{bmatrix}
  1 & 0 & 1 \\
  0 & 1 & 0 \\
  1 & 1 & 0 \\
  \end{bmatrix} \lor \begin{bmatrix}
  1 & 0 & 1 \\
  0 & 1 & 0 \\
  1 & 1 & 0 \\
  \end{bmatrix} = \begin{bmatrix}
  1 & 1 & 1 \\
  1 & 1 & 0 \\
  1 & 1 & 1 \\
  \end{bmatrix}
  \]

  \[
  M_R^{[2]} = M_R \lor M_R^{[2]} = \begin{bmatrix}
  1 & 0 & 1 \\
  0 & 1 & 0 \\
  1 & 1 & 0 \\
  \end{bmatrix} \lor \begin{bmatrix}
  1 & 0 & 1 \\
  0 & 1 & 0 \\
  1 & 1 & 0 \\
  \end{bmatrix} = \begin{bmatrix}
  1 & 1 & 1 \\
  1 & 1 & 0 \\
  1 & 1 & 1 \\
  \end{bmatrix}
  \]
Transitive closure algorithm

• What we did (or rather, could have done):
  – Compute the next matrix $M_i$, where $1 \leq i < n$
  – Do a Boolean join with the previously computed matrix

• For our example:
  – Compute $M_i = M_j \cdot M_j$
  – Join that with $M_j$ to yield $M_j \cdot M_i$
  – Compute $M_i = M_i \cdot M_i$
  – Join that with $M_j \cdot M_i$ from above

Transitive closure algorithm

procedure transitive_closure ($M$: zero-one $n \times n$ matrix)
A := $M_R$
B := A
for $i := 2$ to $n - 1$
begin
  A := A $\cdot M_i$
  B := B $\lor$ A
end { B is the zero-one matrix for $R^*$ }

Transitive closure algorithms

• More efficient algorithms exist, such as Warshall’s algorithm
  – We won’t be studying it in this class
Equivalence vs Partial Order

- Certain combinations of relation properties are very useful
  - Equivalence relations
    - A relation that is reflexive, symmetric and transitive
  - Partial orderings
    - A relation that is reflexive, antisymmetric, and transitive

The difference is whether the relation is symmetric or antisymmetric.

Equivalence relations

- A relation on a set \( A \) is called an *equivalence relation* if it is reflexive, symmetric, and transitive
- Consider relation \( R = \{(a,b) \mid \text{len}(a) = \text{len}(b)\} \)
  - Where \( \text{len}(a) \) means the length of string \( a \)
  - It is reflexive: \( \text{len}(a) = \text{len}(a) \)
  - It is symmetric: if \( \text{len}(a) = \text{len}(b) \), then \( \text{len}(b) = \text{len}(a) \)
  - It is transitive: if \( \text{len}(a) = \text{len}(b) \) and \( \text{len}(b) = \text{len}(c) \), then \( \text{len}(a) = \text{len}(c) \)
  - Thus, \( R \) is an equivalence relation

Equivalence relation example

- Consider the relation \( R = \{(a,b) \mid m \mid a-b\} \)
  - Called "congruence modulo \( m \)"
  - Is it reflexive: \( (a,a) \in R \) means that \( m \mid a-a \)
    - \( a-a = 0 \), which is divisible by \( m \)
  - Is it symmetric: if \( (a,b) \in R \) then \( (b,a) \in R \)
    - \( (a,b) \) means that \( m \mid a-b \)
    - Or that \( km = a-b \). Negating that, we get \( b-a = km \)
    - Thus, \( m \) divides \( b-a \), so \( (b,a) \in R \)
  - Is it transitive: if \( (a,b) \in R \) and \( (b,c) \in R \) then \( (a,c) \in R \)
    - \( (a,b) \) means that \( m \mid a-b \)
    - \( (b,c) \) means that \( m \mid b-c \)
    - Adding these two, we get \( km + km = (a-b) + (b-c) \)
    - Or \( k=km = a-c \)
    - Thus, \( m \) divides \( a-c \), where \( n = k+m \)
  - Thus, congruence modulo \( m \) is an equivalence relation
Sample questions

• Which of these relations on \{0, 1, 2, 3\} are equivalence relations?
  Determine the properties of an equivalence relation that the others lack

a) \{(0,0), (1,1), (2,2), (3,3)\}
   - Has all the properties, thus, is an equivalence relation

b) \{(0,0), (0,2), (2,0), (2,2), (2,3), (3,2), (3,3)\}
   - Not reflexive: (1,1) is missing
   - Not transitive: (0,2) and (2,3) are in the relation, but not (0,3)

c) \{(0,0), (1,1), (1,2), (2,1), (2,2), (3,3)\}
   - Has all the properties, thus, is an equivalence relation

d) \{(0,0), (1,1), (1,3), (2,2), (2,3), (3,1), (3,2), (3,3)\}
   - Not transitive: (1,3) and (3,2) are in the relation, but not (1,2)

e) \{(0,0), (0,2), (1,0), (1,1), (1,2), (2,0), (2,2), (3,3)\}
   - Not symmetric: (1,2) is present, but not (2,1)
   - Not transitive: (2,0) and (0,1) are in the relation, but not (2,1)

Sample questions

• Suppose that \(A\) is a non-empty set, and \(f\) is a function that has \(A\) as its domain. Let \(R\) be the relation on \(A\) consisting of all ordered pairs \((x,y)\) where \(f(x) = f(y)\)
  – Meaning that \(x\) and \(y\) are related if and only if \(x\) and \(y\) have the same image under \(f\).
• Show that \(R\) is an equivalence relation on \(A\)
  - Reflexivity: \(f(x) = f(x)\)
    - True, as given the same input, a function always produces the same output
  - Symmetry: if \(f(x) = f(y)\) then \(f(y) = f(x)\)
    - True, by the definition of equality
  - Transitivity: if \(f(x) = f(y)\) and \(f(y) = f(z)\) then \(f(x) = f(z)\)
    - True, by the definition of equality

Sample questions

• Show that the relation \(R\), consisting of all pairs \((x,y)\) where \(x\) and \(y\) are bit strings of length three or more that agree except perhaps in their first three bits, is an equivalence relation on the set of all bit strings

• Let \(f(x)\) be the bit string formed by deleting the first 3 bits of \(x\), i.e., the last \(n-3\) bits of the bit string \(x\) where \(n\) is the length of the string.

• Thus, we want to show: let \(R\) be the relation on \(A\) consisting of all ordered pairs \((x,y)\) where \(f(x) = f(y)\)

• This has been shown in question 5 on the previous slide
Equivalence classes

• Let $R$ be an equivalence relation on a set $A$. The set of all elements that are related to an element $a$ of $A$ is called the \textit{equivalence class} of $a$.

• The equivalence class of $a$ with respect to $R$ is denoted by $[a]_R$.

• When only one relation is under consideration, the subscript is often deleted, and $[a]$ is used to denote the equivalence class.

• Note that these classes are disjoint!
  – As the equivalence relation is symmetric

More on equivalence classes

• Consider the relation $R = \{ (a,b) \mid a \text{ mod } 2 = b \text{ mod } 2 \}$
  – Thus, all the even numbers are related to each other
  – As are the odd numbers

• The even numbers form an equivalence class
  – As do the odd numbers

• The equivalence class for the even numbers is denoted by $[2]$ (or $[4]$, or $[784]$, etc.)
  – $[2] = \{-4, -2, 0, 2, 4, \ldots\}$
  – $2$ is a representative of its equivalence class

• There are only 2 equivalence classes formed by this equivalence relation

More on equivalence classes

• Consider the relation $R = \{ (a,b) \mid a = b \text{ or } a = -b \}$
  – Thus, every number is related to its additive inverse

• The equivalence class for an integer $a$:
  – $[7] = \{-7, 7\}$
  – $[0] = \{0\}$
  – $[a] = \{a, -a\}$

• There are an infinite number of equivalence classes formed by this equivalence relation
Partitions

• Consider the relation \( R = \{ (a, b) \mid a \mod 2 = b \mod 2 \} \)
• This splits the integers into two equivalence classes: even numbers and odd numbers
• Those two sets together form a partition of the integers
• Formally, a partition of a set \( S \) is a collection of non-empty mutually-disjoint subsets of \( S \) whose union is \( S \)
• In this example, the partition is \( \{ [0], [1] \} \)
  – Or \( \{ \ldots, -3, -1, 1, 3, \ldots \}, \{ \ldots, -2, 0, 2, 4, \ldots \} \)
Partial ordering examples

• Show that $\geq$ is a partial order on the set of integers
  – It is reflexive: $a \geq a$ for all $a \in \mathbb{Z}$
  – It is antisymmetric: if $a \geq b$ then the only way that $b \geq a$ is when $b = a$
  – It is transitive: if $a \geq b$ and $b \geq c$, then $a \geq c$

• Note that $\geq$ is the partial ordering on the set of integers
• $(\mathbb{Z}, \geq)$ is the partially ordered set, or poset

Symbol usage

• The symbol $\preceq$ is used to represent any relation when discussing partial orders
  – Not just the less than or equals to relation
  – Can represent $\leq$, $\geq$, $\subseteq$, etc
  – Thus, $a \prec b$ denotes that $(a, b) \in R$
  – The poset is $(S, \prec)$

• The symbol $\prec$ is used to denote $a \prec b$ but $a \neq b$
  – If $\preceq$ represents $\geq$, then $\prec$ represents $>$

Comparability

• The elements $a$ and $b$ of a poset $(S, \preceq)$ are called comparable if either $a \preceq b$ or $b \preceq a$.
  – Meaning if $(a, b) \in R$ or $(b, a) \in R$
  – It can’t be both because $\preceq$ is antisymmetric
    • Unless $a = b$, of course
  – If neither $a \preceq b$ nor $b \preceq a$, then $a$ and $b$ are incomparable
    • Meaning they are not related to each other

• If all elements in $S$ are comparable, the relation is a total ordering
Comparability examples

• Let \( \leq \) be the “divides” operator \( | \)
• In the poset \((\mathbb{Z}^+, |)\), are the integers 3 and 9 comparable?
  – Yes, as 3 \( | \) 9
• Are 7 and 5 comparable?
  – No, as 7 \( \not| \) 5 and 5 \( \not| \) 7

• Thus, as there are pairs of elements in \( \mathbb{Z}^+ \) that are not comparable, the poset \((\mathbb{Z}^+, |)\) is a partial order

Comparability examples

• Let \( \leq \) be the relation ≤
• In the poset \((\mathbb{Z}^+, \leq)\), are the integers 3 and 9 comparable?
  – Yes, as 3 \( \leq \) 9
• Are 7 and 5 comparable?
  – Yes, as 5 \( \leq \) 7

• As all pairs of elements in \( \mathbb{Z}^+ \) are comparable, the poset \((\mathbb{Z}^+, \leq)\) is a total order
  – a.k.a. totally ordered poset, linear order, chain, etc.

Well-ordered sets

• \((S, \prec)\) is a well-ordered set if:
  – \((S, \prec)\) is a totally ordered poset
  – Every non-empty subset of \( S \) has at least element

• Example: \((\mathbb{Z}, \leq)\)
  – It is a total ordered poset (every element is comparable to every other element)
  – It has no least element
  – Thus, it is not a well-ordered set

• Example: \((S, \prec)\) where \( S = \{ 1, 2, 3, 4, 5 \} \)
  – It is a total ordered poset (every element is comparable to every other element)
  – Has a least element (1)
  – Thus, it is a well-ordered set

• Example: \((\mathbb{Z}^+, \prec)\) is a well-ordered set.
Lexicographic ordering

- Consider two posets: \((S, \prec_1)\) and \((T, \prec_2)\)
- We can order Cartesian products of these two posets via lexicographic ordering
  - Let \(s_1 \in S\) and \(s_2 \in S\)
  - Let \(t_1 \in T\) and \(t_2 \in T\)
  - \((s_1, t_1) \prec (s_2, t_2)\) if either:
    * \(s_1 \prec_1 s_2\)
    * \(s_1 = s_2\) and \(t_1 \prec_2 t_2\)
- Lexicographic ordering is used to order dictionaries.

Lexicographic ordering

- Let \(S\) be the set of word strings (i.e. no spaces)
- Let \(T\) be the set of strings with spaces
- Both the relations are alphabetic sorting
  - We will formalize alphabetic sorting later
- Thus, our posets are: \((S, \prec)\) and \((T, \prec)\)
- Order ("run", "noun: to...") and ("set", "verb: to...")
  - As "run" < "set", the first Cartesian product comes before the "set" one
- Order ("run", "noun: to...") and ("run", "verb: to...")
  - Both the first part of the Cartesian products are equal
  - "noun" is first (alphabetically) than "verb", so it is ordered first

Lexicographic ordering

- We can do this on more than 2-tuples
- \((1,2,3,5) \prec (1,2,4,3)\)
  - When \(\prec\) is \(\leq\)
Lexicographic ordering

• Consider the two strings $a_1a_2a_3\ldots a_m$ and $b_1b_2b_3\ldots b_n$
• Here follows the formal definition for lexicographic ordering of strings

• If $m = n$ (i.e. the strings are equal in length)
  - $(a_1, a_2, a_3, \ldots, a_m) < (b_1, b_2, b_3, \ldots, b_n)$ using the comparisons just discussed
  - Example: "run" < "set"

• If $m \neq n$, then let $t$ be the minimum of $m$ and $n$
  - Then $a_1a_2a_3\ldots a_t$ is less than $b_1b_2b_3\ldots b_t$ if and only if either of the following are true:
    - $(a_1, a_2, a_3, \ldots, a_t) < (b_1, b_2, b_3, \ldots, b_t)$
    - Example: "run" < "sets" ($t = 3$)
    - $(a_1, a_2, a_3, \ldots, a_t) = (b_1, b_2, b_3, \ldots, b_t)$ and $m < n$
    - Example: "run" = "running"

Hasse Diagrams

• Consider the graph for a finite poset $\{1,2,3,4\}$
• When we KNOW it's a poset, we can simplify the graph

Called the Hasse diagram

Hasse Diagram

• For the poset $\{1,2,3,4,6,8,12\}$