Sets & Collections

Lecturer:

Professor Barry O’Sullivan

Office: 2-65, Western Gateway Building

email: *b.osullivan@cs.ucc.ie*

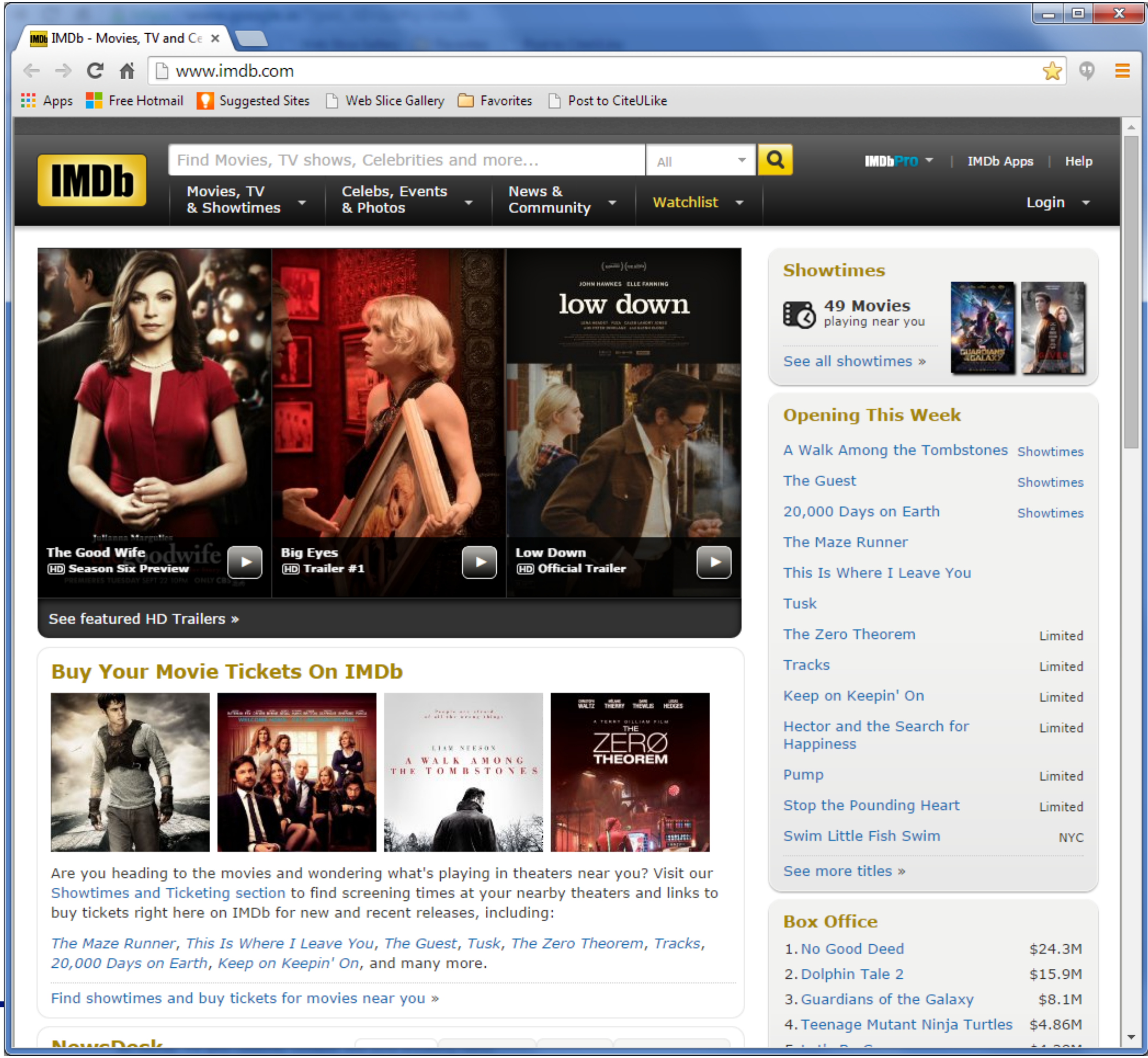
http://osullivan.ucc.ie/teaching/cs1112/

Sets and Collections

Why we need to specify collections of objects

Review of basic set notation

Operations on sets:
Union
Intersection
Difference

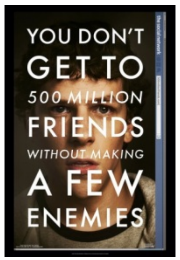

IMDb - Movies, TV and Ce
www.imdb.com
Apps
Free Hotmail
Suggested Sites
Web Slice Gallery
Favorites
Post to CiteULike
IMDb
Find Movies, TV shows, Celebrities and more...
All
IMDbPro
IMDb Apps
Help
Movies, TV & Showtimes
Celebs, Events & Photos
News & Community
Watchlist
Login
low down
The Good Wife
HD Season Six Preview
Big Eyes
HD Trailer #1
Low Down
HD Official Trailer
See featured HD Trailers »
Showtimes
49 Movies
playing near you
See all showtimes »
Opening This Week
A Walk Among the Tombstones Showtimes
The Guest Showtimes
20,000 Days on Earth Showtimes
The Maze Runner
This Is Where I Leave You
Tusk
The Zero Theorem Limited
Tracks Limited
Keep on Keepin' On Limited
Hector and the Search for Happiness Limited
Pump Limited
Stop the Pounding Heart Limited
Swim Little Fish Swim NYC
See more titles »
Buy Your Movie Tickets On IMDb
A WALK AMONG THE TOMBSTONES
THE ZERO THEOREM
Are you heading to the movies and wondering what's playing in theaters near you? Visit our Showtimes and Ticketing section to find screening times at your nearby theaters and links to buy tickets right here on IMDb for new and recent releases, including:
The Maze Runner, This Is Where I Leave You, The Guest, Tusk, The Zero Theorem, Tracks, 20,000 Days on Earth, Keep on Keepin' On, and many more.
Find showtimes and buy tickets for movies near you »
Box Office
1. No Good Deed $24.3M
2. Dolphin Tale 2 $15.9M
3. Guardians of the Galaxy $8.1M
4. Teenage Mutant Ninja Turtles $4.86M

MOVIES

Film: Batman Begins
Director: Nolan
Star1: Bale
Star2: Caine
Year: 2005
Genre: Action

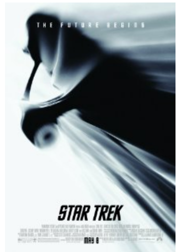
Film: The Hangover
Director: Phillips
Star1: Galifianakis
Star2: Cooper
Year: 2009
Genre: Comedy

Film: The Dark Knight
Director: Nolan
Star1: Bale
Star2: Ledger
Year: 2008
Genre: Action



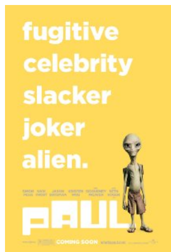
Film: The Social Network
Director: Fincher
Star1: Eisenberg
Star2: Garfield
Year: 2010
Genre: Drama

Film: Inception
Director: Nolan
Star1: DiCaprio
Star2: Page
Year: 2010
Genre: SciFi



Film: Star Trek
Director: Abrams
Star1: Pine
Star2: Pegg
Year: 2009
Genre: SciFi

Film: Avatar
Director: Cameron
Star1: Worthington
Star2: Saldana
Year: 2009
Genre: SciFi

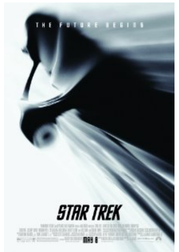


Film: Paul
Director: Mottola
Star1: Pegg
Star2: Frost
Year: 2011
Genre: SciFi

Film	Director	Star1	Star2	Year	Genre
Paul	Mottola	Pegg	Frost	2011	SciFi
Inception	Nolan	DiCaprio	Page	2010	SciFi
The Social Network	Fincher	Eisenberg	Garfield	2010	Drama
Avatar	Cameron	Worthington	Saldana	2009	SciFi
Star Trek	Abrams	Pine	Pegg	2009	SciFi
The Hangover	Phillips	Galifianakis	Cooper	2009	Comedy
The Dark Knight	Nolan	Bale	Ledger	2008	Action
Batman Begins	Nolan	Bale	Caine	2005	Action

SciFi MOVIES

Film: Inception
Director: Nolan
Star1: DiCaprio
Star2: Page
Year: 2010
Genre: SciFi



Film: Star Trek
Director: Abrams
Star1: Pine
Star2: Pegg
Year: 2009
Genre: SciFi

Film: Avatar
Director: Cameron
Star1: Worthington
Star2: Saldana
Year: 2009
Genre: SciFi

Film: Paul
Director: Mottola
Star1: Pegg
Star2: Frost
Year: 2011
Genre: SciFi

MOVIES DIRECTED BY NOLAN

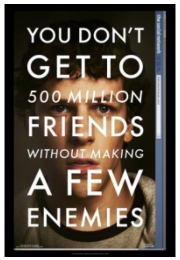
Film:	Batman Begins
Director:	Nolan
Star1:	Bale
Star2:	Caine
Year:	2005
Genre:	Action

Film:	Inception
Director:	Nolan
Star1:	DiCaprio
Star2:	Page
Year:	2010
Genre:	SciFi

Film:	The Dark Knight
Director:	Nolan
Star1:	Bale
Star2:	Ledger
Year:	2008
Genre:	Action

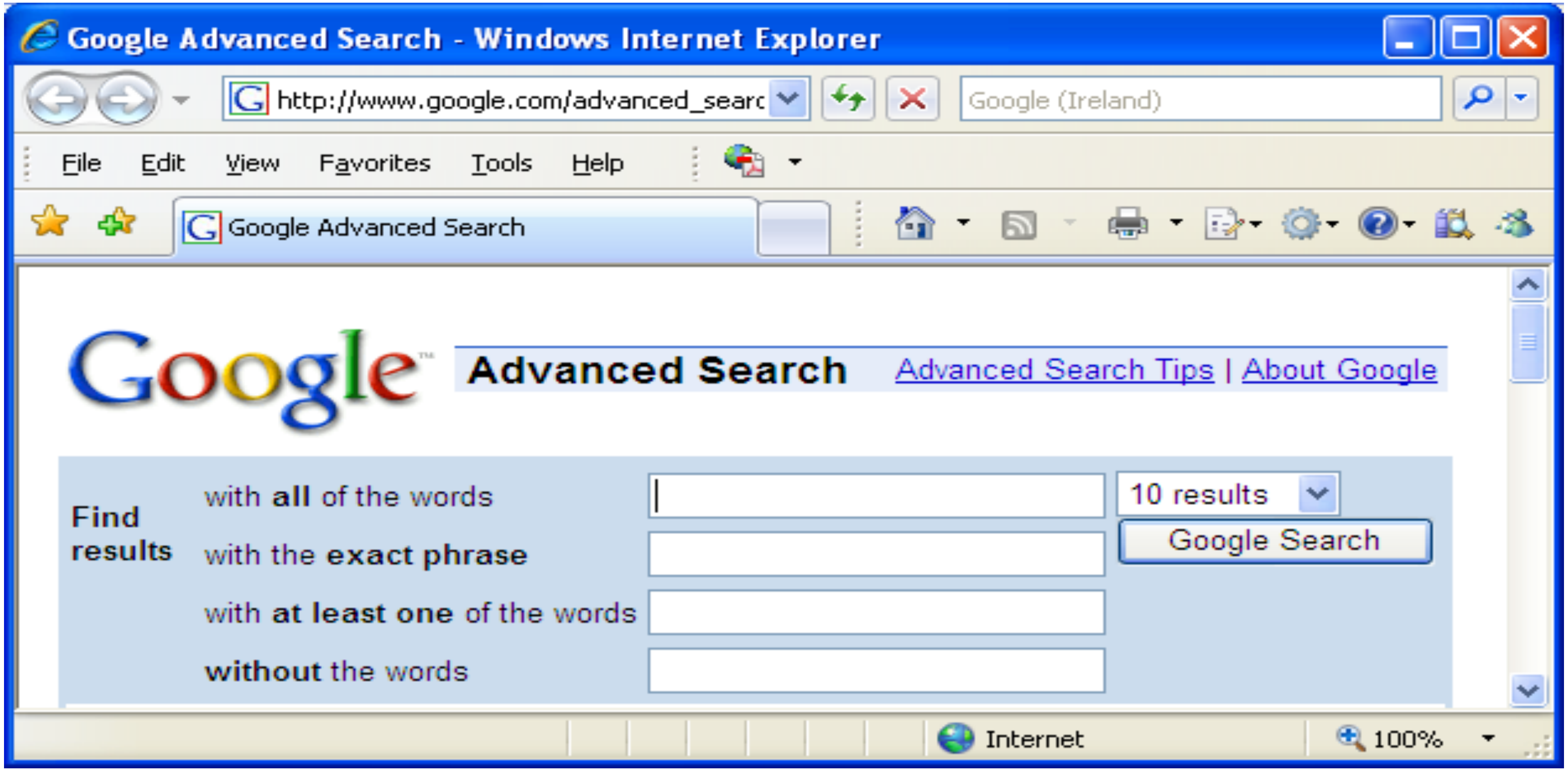
2010 MOVIES

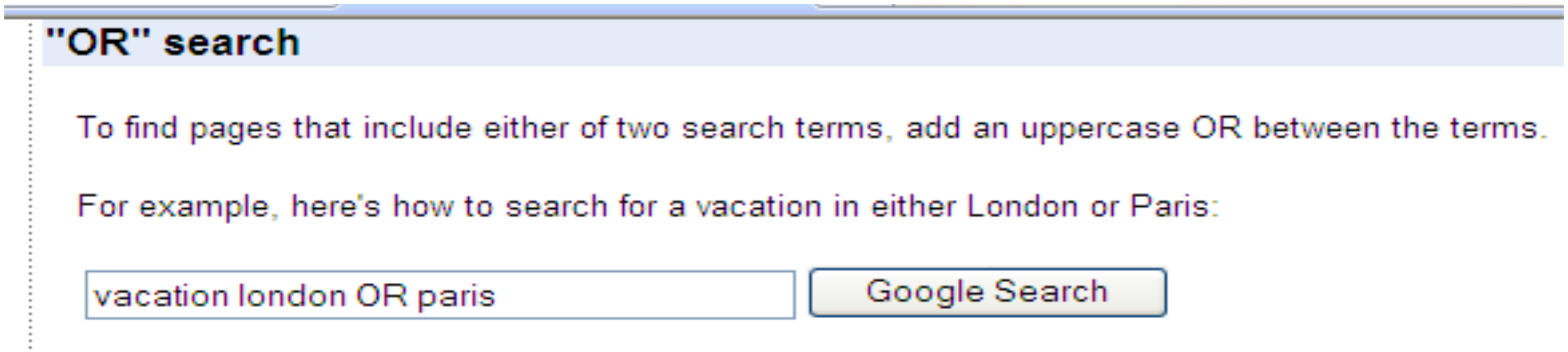
Film:	Inception
Director:	Nolan
Star1:	DiCaprio
Star2:	Page
Year:	2010
Genre:	SciFi



Film:	The Social Network
Director:	Fincher
Star1:	Eisenberg
Star2:	Garfield
Year:	2010
Genre:	Drama


Amazon.co.uk: A First Course in Database Systems: Jeffrey D. Ullman, Jennifer Widom: Books - Windows Internet Explorer
http://www.amazon.co.uk/First-Course-Database-Systems/dp/013600637X/ref=sr_1_2?ie=UTF8&s=books&qid=1222445903&sr=8-2
File Edit View Favorites Tools Help
CS1105 Foundations of Com…
Amazon.co.uk: low prices in …
Amazon.co.uk: A First Co…
Books
Advanced Search
Browse Genres
New & Future Releases
Bestsellers
Paperbacks
Audio Books
Bargain Books
Join Amazon Prime and get unlimited Free One-Day Delivery. Already a member? Sign in.
A FIRST COURSE IN DATABASE SYSTEMS
Jeffrey D. Ullman
Jennifer Widom
A First Course in Database Systems (Hardcover)
by Jeffrey D. Ullman (Author), Jennifer Widom (Author)
No customer reviews yet. Be the first.
RRP: £48.99
Price: £47.49 & this item Delivered FREE in the UK with Super Saver Delivery. See details and conditions
You Save: £1.50 (3%)
In stock.
Dispatched from and sold by Amazon.co.uk. Gift-wrap available.
Only 2 left in stock--order soon (more on the way).
Want guaranteed delivery by Tuesday, September 30? Order it in the next 69 hours and 40 minutes, and choose Express delivery at checkout. See Details
Share your own customer images
Publisher: learn how customers can search inside this book.
26 used & new available from £45.66
Other Editions: RRP: Our Price: Other Offers:
Hardcover 16 used & new from £0.50
Paperback 5 used & new from £13.98
Customers Who Bought This Item Also Bought
Introduction to Algorithms by TH Cormen
(32) £33.10
Data Structures and Algorithm Analysis in Java by Mark Allen Weiss
Data Structures and Algorithms in Java by Michael T. Goodrich
Computer Organization and Design, Revised Printing, Third... by Patterson


Google Advanced Search - Windows Internet Explorer
http://www.google.com/advanced_searc
Google (Ireland)
File Edit View Favorites Tools Help
Google Advanced Search
Google
Advanced Search
Advanced Search Tips | About Google
Find results
with all of the words
with the exact phrase
with at least one of the words
without the words
10 results
Google Search
Internet
100%


"OR" search
To find pages that include either of two search terms, add an uppercase OR between the terms.
For example, here's how to search for a vacation in either London or Paris:
vacation london OR paris
Google Search

Reasoning about collections of objects

- Successful software engineering requires the ability to describe precisely and clearly the intended behaviour of the software. E.g.
  - Finding exactly those students who are entitled to maintenance support
  - Finding all patients who need to return to hospital for a 2^{nd} course of treatment
  - Listing those individuals who can see your profile in a social networking site

- English is too imprecise and ambiguous to do this reliably

- We need a formal notation, and clear rules for reasoning about relationships between collections: set theory

What is a set?

- A set is a collection of things

- The “things” can be anything, physical or abstract:
  - students registered for CS1105 in 2008
  - ID numbers of students registered for CS1105 in 2008
  - The module codes for the specified modules in CK407 1^{st} year programme
  - The objects in my suitcase when I am going on holiday
  - The rules of association of the GAA
  - The 5 most common CDs ordered by people who also purchased the CD you are looking at on amazon.com

- What matters is that we can state clearly and precisely what “things” are in the collection

Specifying a set

- for simple sets, we write down each thing in a list

between "{" and "}" separated by commas

  - {CS1106, CS1110, CS1112, CS1115, CS1117}
  - {banana, apple, orange}
  - {Cork, Kerry, Tipperary, Clare, Limerick, Waterford}
  - {0,1,2,3,4,5,6,7,8,9}

- This style of writing sets is known as set enumeration, or set display, or extensional definition

- Some times, when the pattern is clear, we will write the first few items, three dots, and the last item
  - E.g., we could write the 4th set above as {0,1,2,...,9}

Specifying more complex sets

- In more complex cases, we specify a template, and then a rule, such that everything that matches the template and obeys the rule is in the set (and nothing else is in the set)
- We still write this between “{“ and “}”, and we separate the two parts with “|”
  - {*y* | *y* is a library book with "databases" in the title}
  - {*x* | *x* is a women resident in Co. Cork, registered with the health service, aged 45 or over, who has not yet had a smear test}
  - {*z* | *z* is divisible by 7, *z* > 0, *z* is an integer}
- This is known as an intensional definition

Note: it doesn’t matter which symbols (e.g. x or y) we use in the template, as long as the rule definition is clear

Set membership

- In order to talk about sets, we need more formal notation. Let A be the name of a set, and let x be some thing.
- To say that x is a member of the set A, we write

  $$x \in A$$

- To say that x is not a member of the set A, we write

  $$x \notin A$$

- E.g
  - CS1112 $\in$ {CS1106, CS1110, CS1112, CS1115, CS1117}
  - $3 \in \{ y \mid y \text{ is a single digit} \}$
  - Galway $\notin$ {z | z is a county in Munster}

Set notation and membership: examples

- Which of the following statements are true?

1. apple $\in$ {banana, apple, orange}
2. 12 $\in$ {0,1,2,3,4,5,6,7,8,9}
3. kerry $\in$ {cork,clare,limerick,waterford,tipperary}
4. "Systems Organisation I" $\notin$ {CS1106, CS1110, CS1112, CS1115, CS1117}
5. 6 $\in$ {x | x is an even integer}
6. Clare $\in$ {z | z is a county in Munster}
7. "westlife" $\in$ {z | z is a password consisting of at least 6 characters, one of which must be a digit}
8. "Boole" $\in$ {x | x is the surname of a lecturer on a 1^{st} year Computer Science module in UCC in 2014/15}

The two most important properties of sets

- The order in which you write down elements in an enumeration does not matter
- There is only one copy of each element in a set, so repeated elements in an enumeration are ignored

$\{1,2,3\}$
$\{3,2,1\}$
$\{1,2,3,2,1,2,3,2,1,2,3\}$
$\{1,1,1,1,1,1,1,1,1,1,1,1,2,3\}$
$\{x \mid x$ is an integer such that $0 < x < 4\}$
$\{ z \mid z$ is an integer such that $0 < z < 4\}$
$\{x \mid x$ is the square of 1, or $1 \leq x \leq 3 \}$

are all the same set

Equal sets

- Two sets are equal if they have exactly the same members
- Two sets are not equal if one of them has a member that the other doesn't
- We use the symbols = and ≠ to represent this

$\{1,2,3\} = \{3,2,1,2\}$

$\{1,2,3\} = \{x \mid x \in Z \text{ such that } 0 < x < 4\}$

$\{CS1110, CS1111\} \neq \{CS2200, CS2201\}$

Some important sets

- The empty set is the set with no elements, and is written as { }, or ø
  - For every possible thing x, $x \notin \{\}$ is true
  - For every possible thing x, $x \in \{\}$ is false
- The set Z is the set of integers = {..., -3,-2,-1,0,1,2,3,...}
- The set Q is the set of rationals, or numbers able to expressed as fractions

  $Q = \{n \mid n = x/y, \text{ for some } x \in Z \text{ and some } y \in Z\}$

- The set R is the set of real numbers, which includes numbers like π and the square root of 2

One set A is a subset of another set B if every member of A is also a member of B

We write this as $A \subseteq B$. For example, $\{1,3\} \subseteq \{1,2,3,4\}$

If a set A is not a subset of B, then there must be at least one member of A that is not a member of B, and we write $A \not\subseteq B$. For example, $\{1,5\} \not\subseteq \{1,2,3,4\}$

Examples: which of the following are subsets of
{CS1110, CS1112, CS1115}?

1. {CS1110, CS1115}
2. {CS1110, CS1112, CS1117}
3. {CS1110, CS1112, CS1115}
4. { }

This is important!

"If a set A is not a subset of B, then there must be at least one member of A that is not a member of B"

For any set A, the following two statements are ***always*** true

$$\emptyset \subseteq A$$

and

$$A \subseteq A$$

Another way to think about equality

- We can also define equality between sets in terms of the subset relationship:

  If $A \subseteq B$ and $B \subseteq A$, then $A = B$
  If $A = B$, then $A \subseteq B$ and $B \subseteq A$

Note: sometimes we will write 2-part definitions like this using "if and only if":

$A \subseteq B$ and $B \subseteq A$ if and only if $A = B$

Sometimes, the fact that A is a subset of A doesn't match our informal use of language, so we also have:

A is a proper subset of B if $A \subseteq B$ and $A \neq B$, written $A \subset B$

Example: list all the subsets and proper subsets of
{CS1110, CS1112}

Subsets:

Proper Subsets:

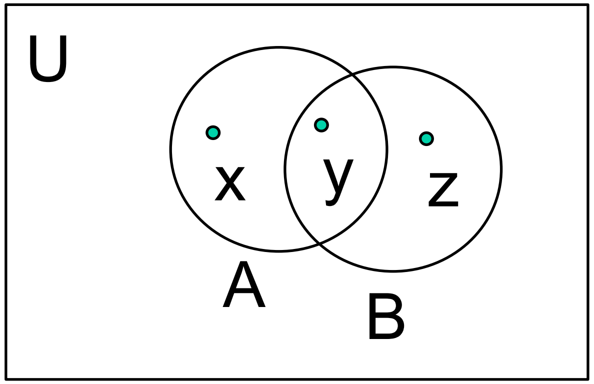
Limiting the scope of our collections

- Sometimes, we will only want to talk about collections of things taken from some clearly defined larger collection
- We call this larger collection the universal set, and if we need to talk about it, we will write it as U
- E.g.
  - when talking about collections of students in UCC (e.g. those registered for CS1112), the universal set is the collection of students who have registered in 2014/15 as a student in UCC
  - When talking about winners of the All-Ireland Senior Hurling Championships, the universal set is the set of counties in Ireland.

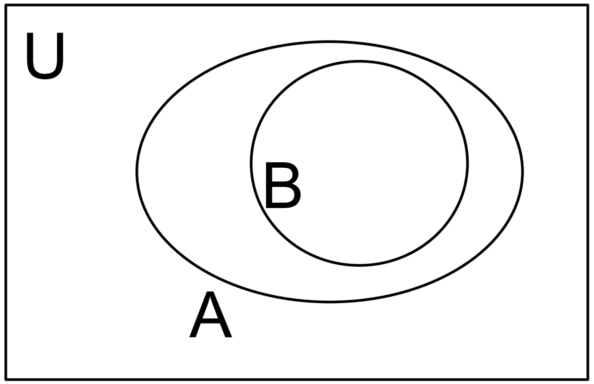
Representing sets using diagrams

Often, you can get a clearer picture of relationships between sets if we draw a diagram.

Draw the universal set as a large rectangle. Draw individual sets as circles (or sometimes arbitrary enclosed shapes to make the drawing easier). Draw individual members as points inside the shape representing the set (or just omit them).



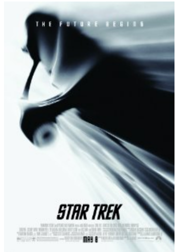
$A = \{x,y\}$ $B = \{y,z\}$

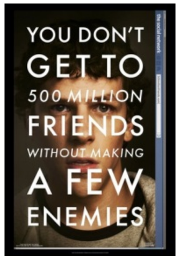


$B \subseteq A$

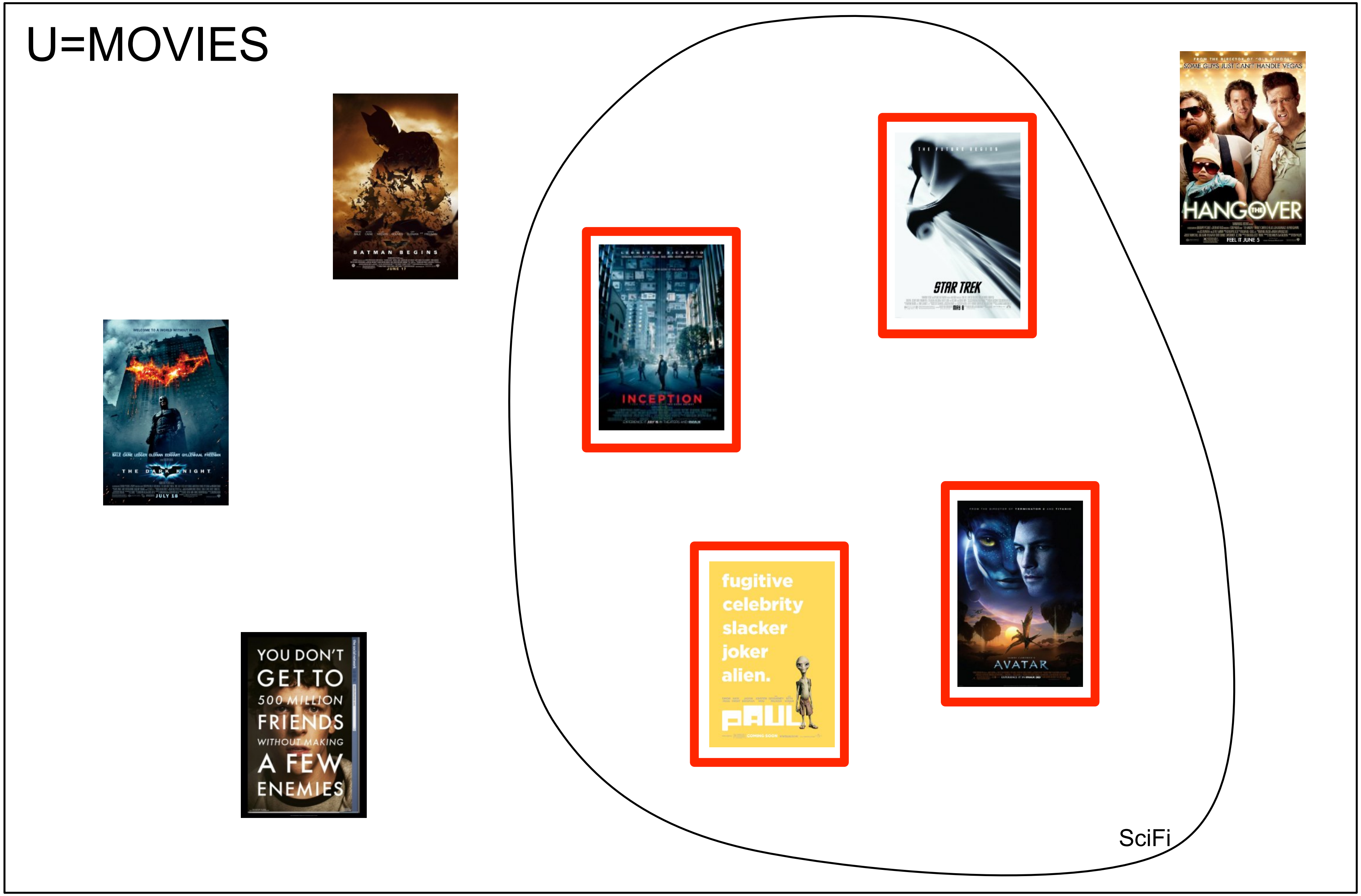
Our (limited) movie database

U=MOVIES





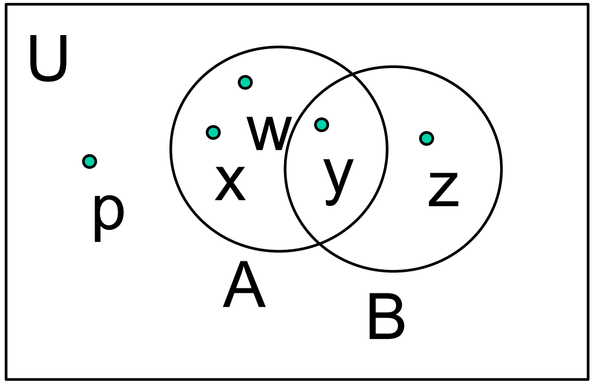
SciFi movies



Nolan movies

The set union of two sets is a new set consisting of every member of the two sets

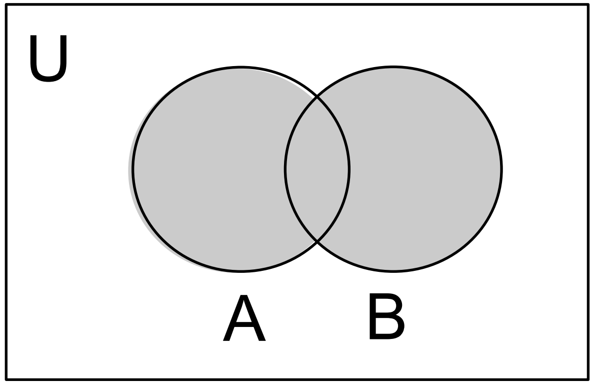
We write this as $A \cup B$



$A = \{x,y,w\}$ $B = \{y,z\}$

$A \cup B = \{x,w,y,z\}$

In general:

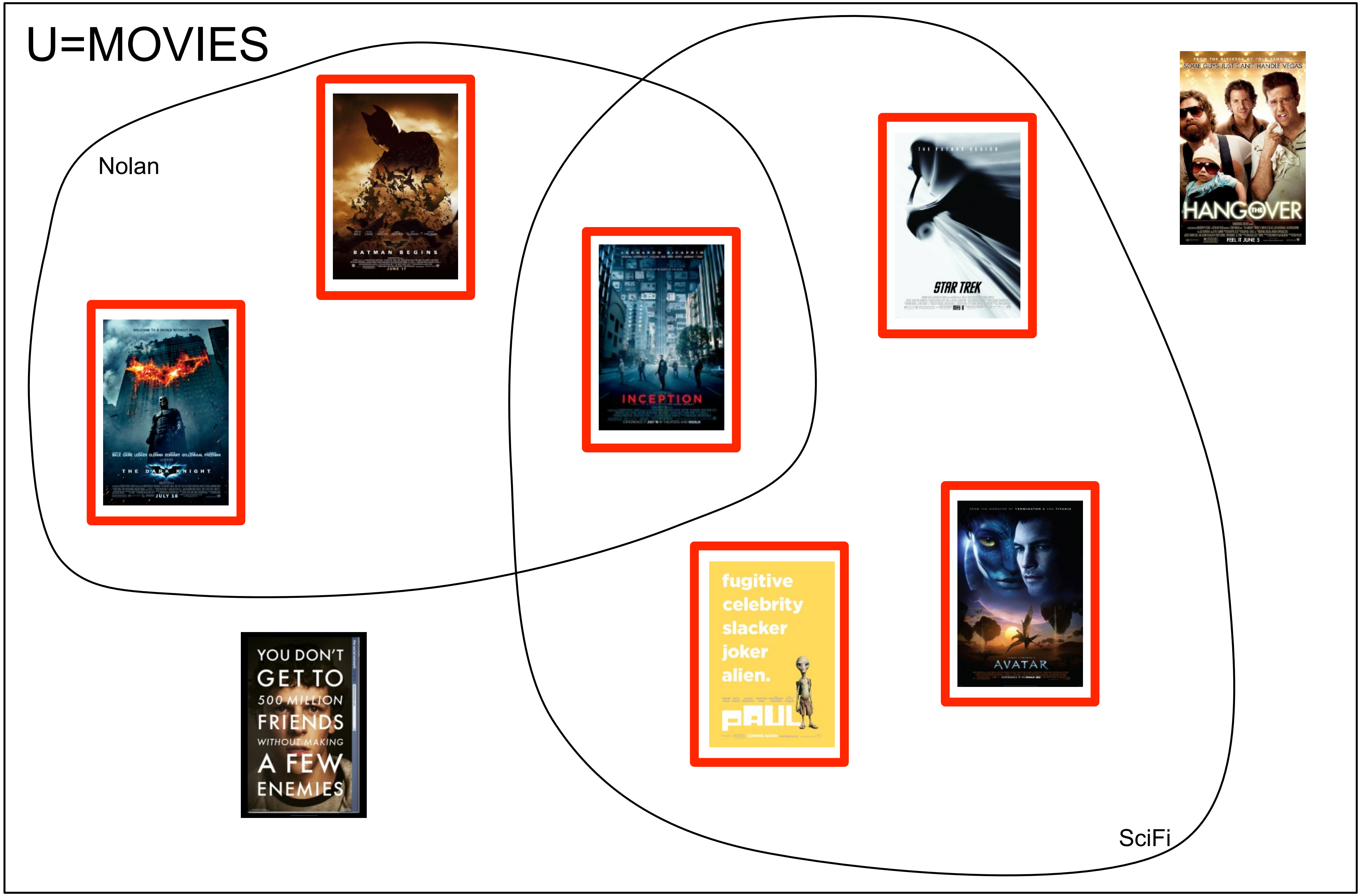


$A \cup B$

For any element x of U,

$x \in A \cup B$ if and only if $x \in A$ or $x \in B$

Nolan ∪ SciFi


U=MOVIES
Nolan
SciFi
BATMAN BEGINS
THE DARK KNIGHT
INCEPTION
STAR TREK
HANGOVER
AVATAR
fugitive
celebrity
slacker
joker
alien.
PAUL
YOU DON'T
GET TO
500 MILLION
FRIENDS
WITHOUT MAKING
A FEW
ENEMIES

Examples:

<u>$A \cup B$?</u>

A={1,3,5}, B = {2,4,6}

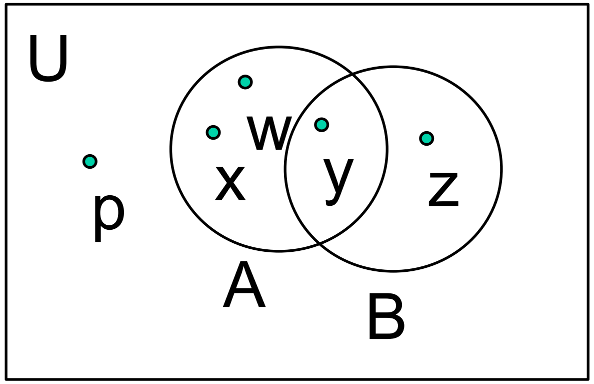
A = {e,g,b,d,f}, B = {f,a,c,e}

A = {CS1101, CS1105,CS1106, CS1107},
B = {CS1101, CS1105}

A= { }, B = {apple, orange, banana}

The set intersection of two sets is a new set consisting of each member that appears in both sets

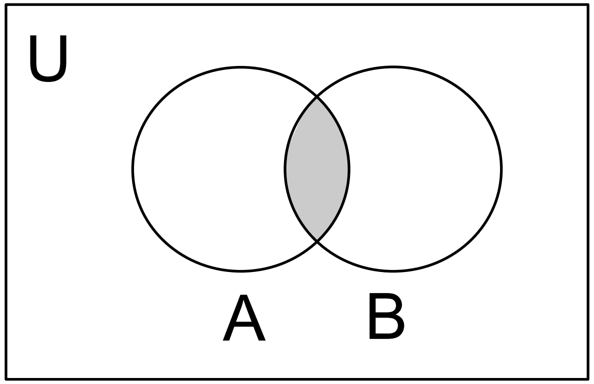
We write this as $A \cap B$



$A = \{x,y,w\}$ $B = \{y,z\}$

$A \cap B = \{y\}$

In general:

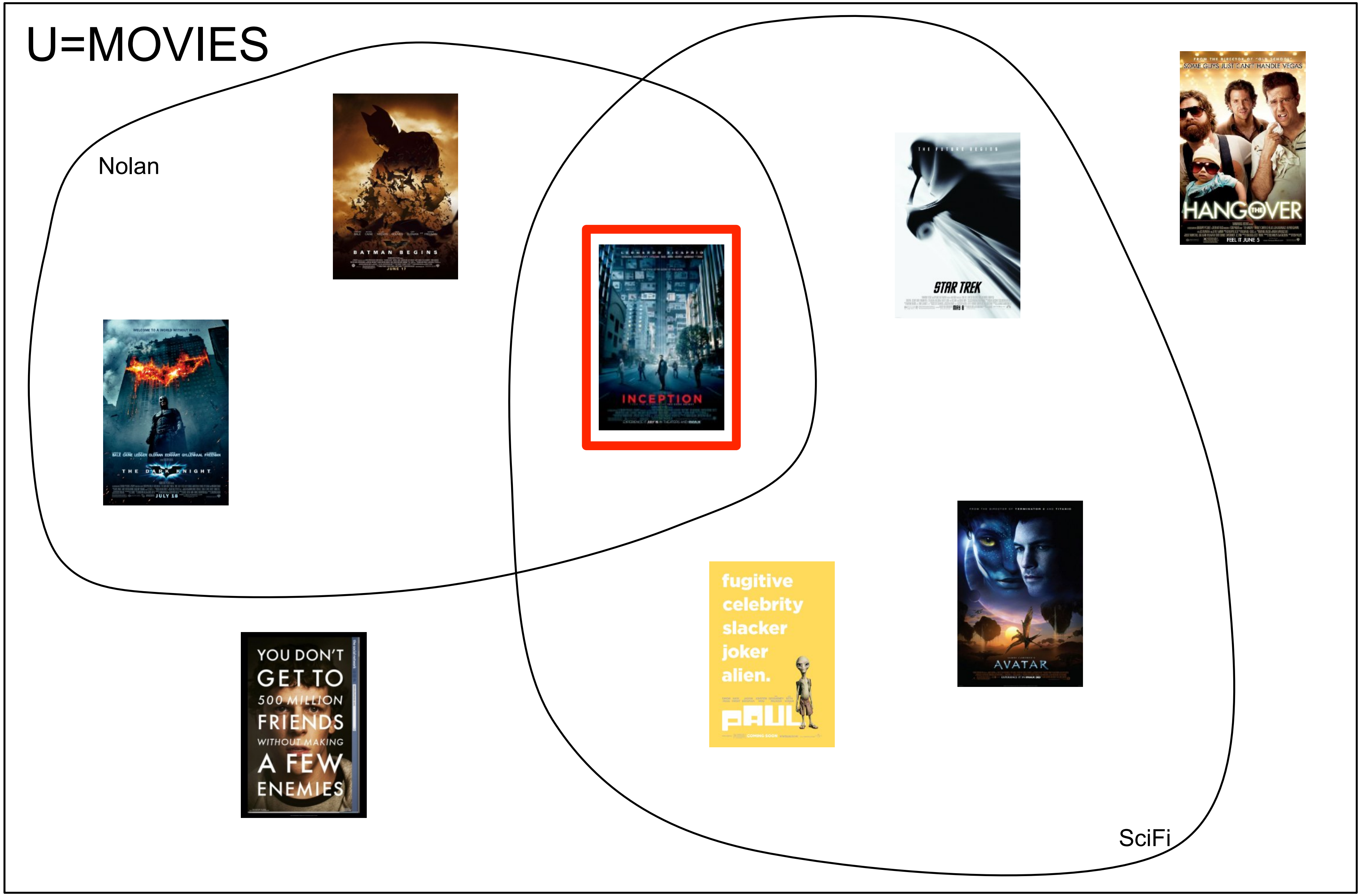


$A \cap B$

For any element x of U,

$x \in A \cap B$ if and only if $x \in A$ and $x \in B$

Nolan ∩ SciFi


U=MOVIES
Nolan
SciFi
BATMAN BEGINS
THE DARK KNIGHT
JULY 18
INCEPTION
STAR TREK
SOME GUYS JUST CAN'T HANDLE VEGAS
HANGOVER
YOU DON'T GET TO 500 MILLION FRIENDS WITHOUT MAKING A FEW ENEMIES
fugitive celebrity slacker joker alien.
PAUL
AVATAR

Examples:

<u>$A \cap B$?</u>

A={1,2,3}, B = {1,3,5}

A = {e,g,b,d,f}, B = {f,a,c,e}

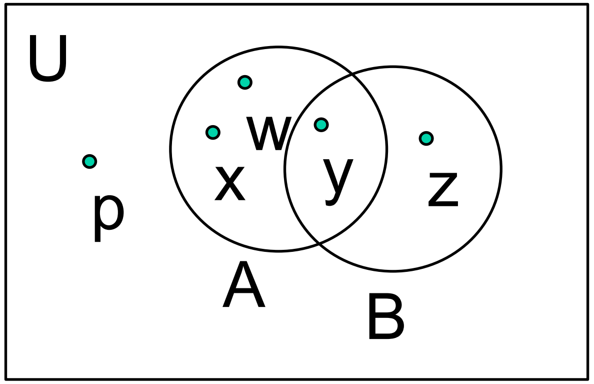
A = {CS1101, CS1105,CS1106, CS1107},
B = {CS1101, CS1105}

A= { }, B = {apple, orange, banana}

A = {cork, kerry}, B = {cork, kerry}

The set difference of two sets is a new set consisting of each element of the first that is not also an element of the second

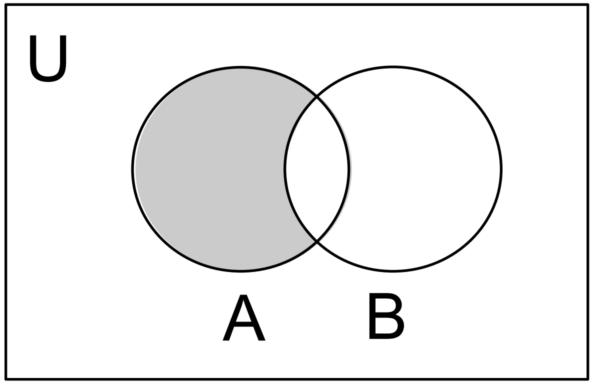
We write this as $A \setminus B$ or sometimes $A - B$



$A = \{x,y,w\}$ $B = \{y,z\}$

$A \setminus B = \{x,w\}$

In general:

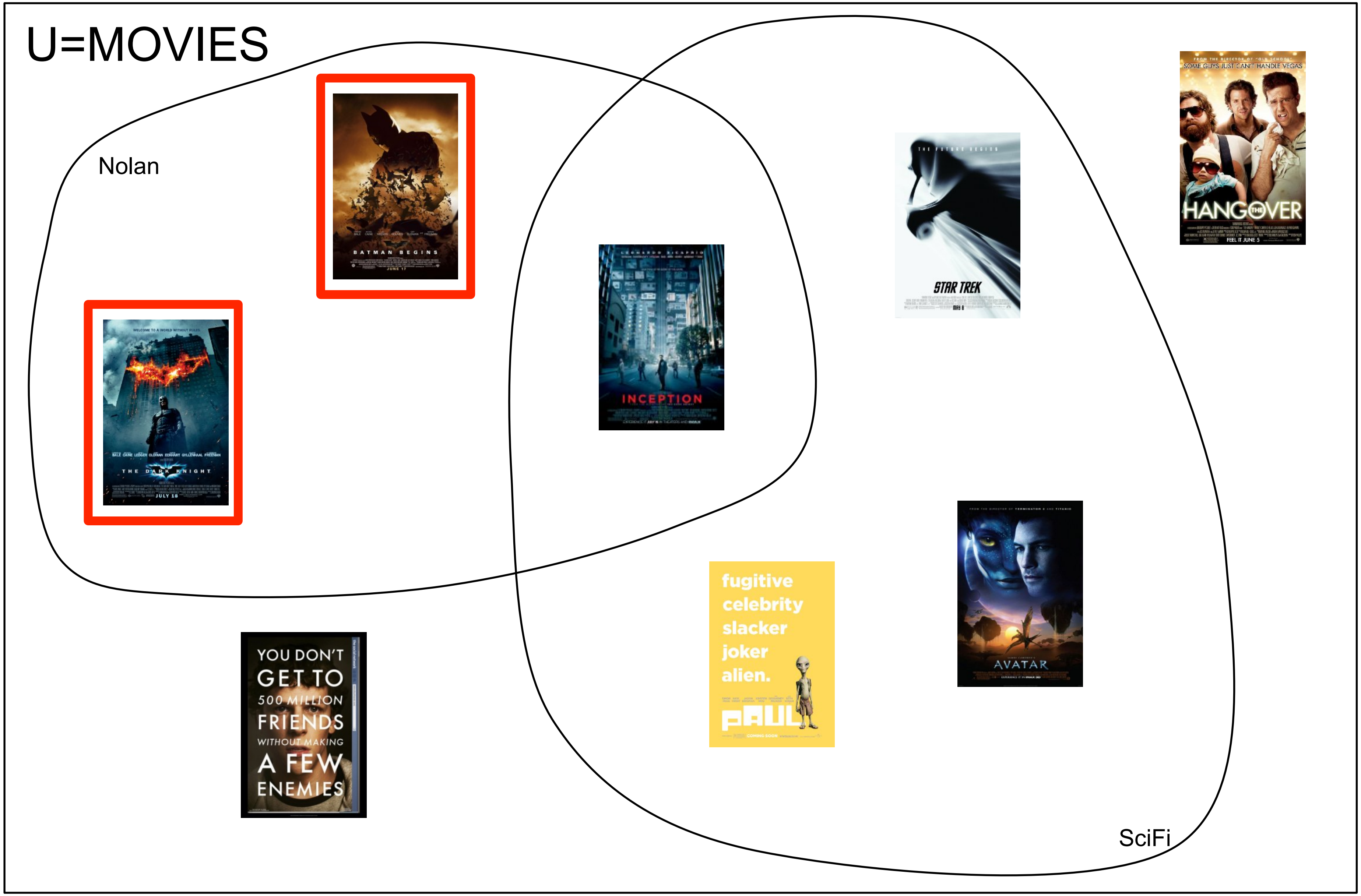


$A \setminus B$

For any element x of U,

$x \in A \setminus B$ if and only if $x \in A$ and $x \notin B$

Nolan \ SciFi


U=MOVIES
Nolan
BATMAN BEGINS
THE DARK KNIGHT
INCEPTION
STAR TREK
SOME GUYS JUST CAN'T HANDLE VEGAS
THE HANGOVER
YOU DON'T GET TO 500 MILLION FRIENDS WITHOUT MAKING A FEW ENEMIES
fugitive
celebrity
slacker
joker
alien.
PAUL
AVATAR
SciFi

Examples:

<u>A \ B ?</u>

A={1,2,3}, B = {1,3,5}

A = {e,g,b,d,f}, B = {f,a,c,e}

A = {CS1101, CS1105,CS1106, CS1107},
B = {CS1101, CS1105}

A= { }, B = {apple, orange, banana}

A = {cork, kerry}, B = {cork, kerry}

A = {cork, kerry}, B = { }

Next lecture ...

Set complement

Cardinality

Sets containing other sets

Power set
Partition

Laws of set operations

Cartesian product