DISCRETE MATHS TO REMEMBER

BASIC PROPERTIES OF NUMBERS

- Types of numbers:
  - Natural numbers: $\mathbb{N} - 0, 1, 2, 3..$, Integers: $\mathbb{Z} - ... - 2, -1, 0, 1, 2...$, Rational numbers $\mathbb{Q} - ... - \frac{1}{2}..., 0, ..., \frac{1}{2}, \frac{a}{b}...$, Real numbers $\mathbb{R}$ is the set of all numbers (rational and irrational), $\mathbb{C}$ is the two dimensional set of all complex numbers.
  - Primes - $2, 3, 5, 7..$ Notice the set does not contain 1 or 0.
- Converting from decimal to binary: Repeatedly divide by 2, keeping track of the remainders. Stop when you get to zero. The binary number will start with its first digit as the last remainder of repeated divisions. For octal and hexadecimal do exactly the same only dividing by 8 and 16 respectively instead.
- Binary arithmetic - just add and multiply like you would normally, carrying numbers correctly.
- Converting to and from binary and octal and hexadecimal. Groups of 3 binary numbers are mapped to octal and groups of 4 are mapped to hexadecimal. Always starting from the right and ending at the left.
- Two's compliment. Switch each digit and then add 1. Remember that the first digit on the left is negative, so $1111_2 = -1_{10}$ for example.
- Modular numbers. $x = y \mod n$ if and only if $n|(x - y)$, i.e. $x - y$ is divisible by n. From this you can derive that $x = an + y$, where $a \in \mathbb{Z}$. Use multiplication tables if you get lost.
- Greatest common divisor(gcd): c is the greatest common divisor of x and y, if it divides both x and y and any other common divisor, say d, then c is divisible by d $(d|c)$.
- Least common multiple (lcm). c is the least common multiple of x and y, if c is a multiple of both x and y and any other multiple, say d, then d is divisible by c $(c|d)$.
- Finding gcd using Euclid's Algorithm. Divide x by y $(x > y)$ recording the remainder (r1). Then divide y by the remainder (r1), recording the next remainder (r2). Then divide r1 by r2 and keep repeating the process until you reach a zero remainder. The gcd will be the last non-zero remainder.
- Finding gcd and lcm generally: Write x and y (these are just two random numbers) as products of primes. So e.g. $x = p_1^{a_1} \times p_2^{a_2} \times ... \times p_n^{a_n}$ and $y = p_1^{b_1} \times p_2^{b_2} \times ... \times p_n^{b_n}$, then,
  - $gcd(x, y) = p_1^{min(a_1,b_1)} \times p_2^{min(a_2,b_2)} \times p_n^{min(a_n,b_n)}$

– $lcm(x,y) = p_1^{max(a_1,b_1)} \times p_2^{max(a_2,b_2)} \times p_n^{max(a_n,b_n)}$

- Fundamental Theorem of Arithmetic. Every **natural** number greater than 1 can be written as a **unique** product of primes. Proof: Suppose there are some natural numbers > 1, which cannot be written as a product of primes. Let k be the smallest such number. Now, k cannot be prime. Therefore, $k = x \times y$, where $x, y < k$. However, since $x, y < k$ then they must be products of primes. So, therefore, k itself is a product of primes. This is a contradiction. Uniqueness follows from 1 not being a prime.

- **No largest prime**. Suppose that $P_k$is the largest prime. Then all primes can be written as $P_1...P_k$. Let $x = (P_1 \cdot P_2...P_k) + 1$. By the fundamental theorem of arithmetic (FTA) x must be a product of primes, so $P_j|x$, where P_j is one of the primes in the list of all primes, for $1 \leq j \leq k$. Therefore, $x = (P_1 \cdot P_2...P_k) + 1 = P_jL$, where L is any natural number greater than 1. If we then divide x by P_j we find that x has a remainder of 1 and a remainder of 0. This is a contradiction, hence there is no largest prime.

Modular numbers - **solving systems of equations**
e.g.

$$x = 2 \bmod 3 \tag{1}$$
$$x = 3 \bmod 5 \tag{2}$$
$$x = 2 \bmod 7 \tag{3}$$

This means:

$$x = 2, 5, 8, 11, 14, 17, 20, \underline{23} \tag{4}$$
$$x = 3, 8, 13, 18, \underline{23} \tag{5}$$
$$x = 2, 9, 16, \underline{23} \tag{6}$$

To find modular simply find the least common multiple of the given modular numbers, in this case $(3 \times 5 \times 7) = 105$. So the solution is $x = 23 \bmod 105$.

To solve analytically, do the following:

$$x = 3a + 2 \tag{7}$$
$$3a + 2 = 3 \bmod 5 \qquad (= 5k + 3) \tag{8}$$
$$3a = 1 \bmod 5 \qquad (= 5k + 1) \tag{9}$$

The next step is tricky, you have to find the values of x which will satisfy the equation and then plug that value in as the new remainder, like this:

$$a = 1 \qquad 3 = 3 \bmod 5 \quad (\neq 1 \bmod 5) \tag{10}$$
$$a = 2 \qquad 6 = 1 \bmod 5 \quad (yay) \tag{11}$$

So,

$$a = 2 \bmod 5 \tag{12}$$
$$a = 5b + 2 \tag{13}$$
$$x = 3(5b + 2) + 2 = 15b + 8 \tag{14}$$

Then repeat steps 10 to 17 with the last equation for x. You should get $x = 105c + 23$. Set c to 0 to get final value for x.

SET THEORY

- Set difference := $A - B = A \cap B^C$
- Complement := $A^C = U - A$
- $A \subset B = A \subseteq B, A \neq B$
- Properties:
  - Basic ones like: De Morgans, Commutative, Associative, Distributive, Complement
  - Idempotent := $A \cup A = A$ and $A \cap A = A$
  - Excluded Middle := $A \cup A^C = U$ and $A \cap A^C = \emptyset$
  - Elimination := $A \cup \emptyset = A$, $A \cap \emptyset = \emptyset$, $A \cup U = U$, $A \cap U = A$.
- Use truth tables if you are analyzing membership functions, or trying to simplify sets
- The partition of a set, is a family of subsets whose union is the set and their intersections is the empty set, i.e. they do not intersect.
- Cardinality is the number of elements in a set. You can use it to show that $|A \cup B| = |A| + |B| - |A \cap B|$
- Power set. This is a set containing all possible sub-sets of a set. E.g. if $A = \{a, b, c\}$, then $P(A) = \{\{a, b, c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a\}, \{b\}, \{c\}, \emptyset\}$. The cardinality of the power set, $|P(A)|$ is equal to $2^{|A|}$. This is because there are 2 possible outcomes whether an element should be kept in a sub-set or not, for n elements. If $A_1 ... A_n$ are partitions of the set B, then the power set of B, then the number of elements in $|P(B)| = |P(A_1)| \times ... |P(A_n)|$

Set notation

$$\text{e.g. } C = \{x : x = 1 \text{ or } x - 3 \in C\} \tag{15}$$

C is a recursive set, so build up from nothing and check conditions. If $U = \{1..10\}$ then plug in U into $x - 3$. Only at $x = 4$ does x satisfy the condition, same with $x = \{7, 10\}$ so finally, $C = \{1, 4, 7, 10\}$. Another example is the set of all natural numbers (including zero), which can be written like this:

$\{0\} \cup \{x+1 : x \in N\}$.

RELATIONS

- A relation is a set of ordered pairs.
- The cross product of two sets is the relation $A \times B = \{(x, y) : x \in A, y \in B\}$.
- Hasse Diagrams represent partial orderings. Remember to remove all reflexive and transitive arrows.
- Incomparable Elements are elements of a set that do not belong to same subset?
- Maximal Elements are the largest elements in the subsets of a partial ordering.
- Minimal Elements are the smallest elements in the subsets of a partial ordering.
- sup C is the least upper bound of a subset (this case set C) of a set.
- inf C is the greatest lower bound of a subset of a set.

Example: $[5, 7]$ - max element = 7, min element = 5. $(5, 7)$ - max and min element not possible. sup(5,7) = 7, inf(5,7) = 5.

Equivalence Relations
To have equivalence check:

- Reflexivity := $R(x, x)$
- Symmetry := $R(x, y) = R(y, x)$
- Transitivity := If $R(x, y)$ and $R(y, z)$ then $R(x, z)$ must hold.
- Equivalent classes put together sets of equivalent elements.
- Equivalent classes form a partition of the set that is being acted upon by a relation.

Partial Orderings To have partial ordering, check:

- Reflexivity := $R(x, x)$ e.g with sub-sets you get $(a_1 \leq a_1)$ and $(b_1 \leq b_1)$ which is true.
- Anti-Symmetry := $R(x, y)$ and $R(y, x)$ then $x = y$ must hold.
- Transitivity := $R(x, y)$ and $R(y, z)$ then $R(x, z)$ must hold.

- Examples of partial orderings are $\leq$ and $\subseteq$, often notated as $(A, \leq)$

FUNCTIONS

- $f : X \to Y$ is a special kind of relation on $A \times B$ which satisfies the following two properties:
  - 1. $\forall x \in X$ there exists $y \in Y$ such that $(x, y) \in f$
  - 2. If $(x, y) \in fthen(x, z) \notin f$ for $z \neq y$
  - $(x, y) \in f$ is the same as $y = f(x)$
  - A partial function only satisfies property 2.
- The domain of f is $dom(f) = X$ and the range of f is $ran(f) = \{f(x) : x \in X\}$.
- f is a surjective function (or Surjection) if $ran(f) = Y$. Set Y is completely filled.
- f is an injective function (or injection) if $\forall a, b \in X, f(a) = f(b)$ which implies that $a = b$. The set X is completely filled.
- Bijective means it is injective and surjective.

PROBABILITY AND INFORMATION

- Probability measure is a function $P : 2^S \to [0, 1]$, where 2^S is the power-set of the sample space S ($P(S)$) satisfying the following conditions:
  - $P(S) = 1$ and $P(\emptyset) = 0$
  - (Additivity) If $A \cap B = \emptyset$ then $P(A \cup B) = P(A) + P(B)$
- Probability distribution - P(s) defined across all elements $s \in S$ is called a probability distribution on S.
- Random Variable. A random variable $X : S \to \mathbb{R}$ is a real valued function defined on S.
- The probability distribution on S defines the probability distribution on X according to: $P(X = x) = P(\{s \in S : X(s) = x\})$
- Shannon's Entropy Measure $H_n(\overrightarrow{p}) = -\sum_{i=1}^{n} p_i log_2(p_i)$
- Unit of entropy is the bit.
- Bayes Rule $:= P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A)+P(B|A^c)P(A^C)}$
- Max entropy is found simply by $H_n(\frac{1}{n}, ...\frac{1}{n}) = log_2(n)$ or by solving $\frac{dH_n}{dp} = 0$

- Information is max entropy minus information provided? $I_n(K) = log_2(n) - log_2(?)$?
- Joint Probability Distributions
  - Two random variables are independent if $P(X = x, Y = y) = P(X = x) \times P(Y = y)$
  - Marginal Distributions are $P(X = x) = \sum_y P(X = x, Y = y)$ and $P(Y = y) = \sum_x P(X = x, Y = y)$ - basically just summing rows or columns of a joint distribution table.
- Joint Entropy
  - $H(X, Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} -p_{i,j} log_2(p_i, j)$ which basically says to calculate the entropy of every combined probability in a table.
  - If random variables X and Y are independent then $H(X, Y) = H(X) + H(Y)$
- Conditional Entropy
  - $H(X_1|Y = 1) = \frac{P(X_1 \cap Y=1)}{P(Y=1)}$ which in other words is the probability from the intersection of $X_1 and Y_1$ divided by the marginal probability $P(Y = 1)$. To find $H(X|Y = y)$ just apply the entropy function to the found probabilities X in row $Y = y$.
  - $H(X|Y) = \sum_{j=1}^{m} P(y_j) H(X|Y = y_j)$ which means you need to sum the following $P(Y = y_1)H(X|Y = y_1) + P(Y = y_2)H(X|Y = y_2)$ ... until basically $y_n = |Y|$.
  - Useful thing to know: $H(X, Y) = H(Y, X) = H(Y) + H(X|Y) = H(X) + H(Y|X)$ (symmetry)
  - Mutual Information $I(X; Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$ (quantifies information about X contained in Y.

LOGIC AND PROOF

Proof by contradiction

- **For divisibility of numbers**. E.g. If $5|x^2$ then $5|x$. Suppose x is not divisible by 5, then $x = 5k + r$, where $r = 1, 2, 3, 4$. Then $x^2 = (5k + r)^2 = 25k^2 + 10kr + r^2$. Since r^2 is not divisible by 5, this is a contradiction and hence $5|k$.

Proof by induction

- Limit Case
- Inductive step
- Write "by inductive hypothesis" while in Inductive Step process.

Contrapositive proof

- "If A then B" is equivalent to "If not A then not B"

Proof by resolution

- Convert sentences to CNF (conjunctive normal form)
- Remember to take the $\neg$ of what is implied, i.e. after the $\vdash$ (implies) symbol.
- Set up sentences in a list and try to find the empty set.

KEY Equivalences

$$\theta \rightarrow \varphi = \neg\theta \vee \varphi \tag{16}$$
$$\theta \leftrightarrow \varphi = (\theta \rightarrow \varphi) \wedge (\varphi \rightarrow \theta) \tag{17}$$

- Inference rule. An excepted step in an argument to show that a new statement follows from existing ones.
- Modus ponens (example of an inference rule). Simply, **if A then B**. Also written as $\frac{A}{B}$
- Syllogism. (Also an inference rule) **All X are Y. Z is a X then Z is a Y**.
- Proof by induction (inference rule for natural numbers): 1. P(a) holds 2. For all n, if P(n) holds then P(n + 1) holds. Therefore 3. For every $n \geq a$, P(n) holds.
- Proof by contradiction
  - Diophantine equation. There are strictly no positive integer (Natural numbers without zero) solutions to the Diophantine equation $x^2 + y^2 = 1$.
  - $\sqrt{2}$ is irrational.
- Contrapositive proof - **If A then B** is the same as **If not B then not A**.
  - Example: Prove that if n is a positive integer not equal to 2 and n is a prime then n is odd. To do this, show instead that, if n is even then either $n = 2$ or n is not prime. Suppose n is even then $n = 2k$ for some positive integer $k < n$. If $k = 1$ then $n = 2$. If $k > 1$ then n is divisible both by 2 and k, hence is cannot be prime.
- Propositional logic.
  - Declarative sentence is a sentence that is either true or false.

- Rules: All propositional variables and declarative sentences are grammatically correct sentences. Basically propositional variables are declarative sentences.
- Brackets are added in the following order: $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$
- Valuation is simply the 'function' that gives something true a value of 1 and something false a value of 0.
- conjunction ($\wedge$) truth table top to bottom: 1000.
- disjunction ($\vee$) 1110. Remember that false and true is true.
- implication ($\rightarrow$) 1011. Remember false can imply true.
- If and only if ($\leftrightarrow$) 1001.
- Idempotence - $\theta \wedge \theta \equiv \theta$. De Morgan's law, Commutativity, Associativity, Distributivity all apply. (Abelian group and field hihi).
- Contrapositive $\theta \rightarrow \psi \equiv \neg\psi \rightarrow \neg\theta$
- Tautology. A sentence which is true in every valuation.(law of excluded middle) e.g. $(\theta \vee \neg\theta)$. Tautology elimination - $\theta\wedge$ tautology $= \theta$)
- Contradiction. A sentence which is false in every valuation. (Law of non-contradiction) e.g. $(\theta \wedge \neg\theta)$ Contradiction elimination - $\theta\underline{\vee}$ contradiction $= \theta$
- Follows from (entailment) $\models$, means that two sentences have the same truth valuation for every column of one row in the truth table, i.e. getting 1 1 1 1 = 1 in one of the rows. It is a *semantic* operator (relies on truth tables) operator.
- The proof for resolution symbol ($\vdash$) is a *proof theory* operator (based on axioms and rules of inference).
- Soundness means for any sentence θ and ψ, if $\theta \vdash \psi$ then $\theta \models \psi$ (obvious when you think about it)
- Completeness. If $\theta \models \psi$ then $\theta \vdash \psi$. (Only sometimes the case).

- Predicate logic
  - Predicates PL: P(x) means x has property P. P(x,y) means x and y have property P.
  - Functions FNL: f(x) takes x and maps it to something else. e..g f(x) = 'father of x'.
  - Constants CL: 0,1, Bob etc.., Variables, x, y, z.
  - Terms: possible instantiations of predicates, basically members of the set that contains all possible cases of the predicate. e.g. FNL = father, mother.
  - Atomic formulae - basically a predicate with terms instead of variables - $P(t_1, t_2.., t_n)$.

- Well-Formed Formulae. Basically just a set of well formed predicate logic.
- $\forall$ means 'for all'. $\exists$ means 'there exists'.
- Implies ($\rightarrow$) 'If A then B'.
- Equivalent ($\leftrightarrow$) 'If A if and only if B'.

INTRODUCTION TO ALGEBRA

- The bijective function $f : A \rightarrow A$ is called a permutation A. This is because, if say $|A|$ = n, then there are $n!$ permutations that the function f can map set A onto set A.
- Permutations are a **Group** (called symmetric group) - so they have closure, associativity, an identity and an inverse. Proof for associativity is as follows: For $x \in A$, $f \circ (g \circ h)(x) = f(g \circ h(x)) = f(g(h(x))) = (f \circ g)(h(x)) = (f \circ g) \circ h(x)$. Note, $f \circ g = f(g(x))$ means apply permutation g first and then apply permutation f.
- Quaternions are another example of a group.

A group $(\mathbf{G}, \cdot)$, is a set **G** and a binary operation f(x,y) or $x \cdot y$, which looks something like this: $(\{1, 2, 3, 4\}, \times)$. It must satisfy the following four conditions:

- Closure := $x, y \in \mathbf{G}$ then $x \cdot y \in \mathbf{G}$
- Associativity := $\forall x, y, z \in \mathbf{G}$ then $x \cdot (y \cdot z) = (x \cdot y) \cdot z$. Note: Binary operations multiplication and addition are always associative.
- There must be an identity e ($\exists e \in G$), such that $x \cdot e = e \cdot x = x$
- Each element in **G** has an inverse that is also in **G** := $x \cdot x^{-1} = e$
- Abelian Group is **group** with the additional property of being commutative - $\forall x, y \in G, x \cdot y = y \cdot x$
- Proof that an element n has an inverse n^{-1} under multiplication mod p for which $n \times n^{-1} = 1 \bmod p$ if and only if n is relatively prime (i.e. $gcd(n, p) = 1$): $n \times n^{-1} = 1 \bmod p = p(a) + 1$ for some a. Suppose $gcd(n, p) > 1$ and there exists a $k > 1$ such that $k|n$ and $k|p$. Thus, since $1 - n \times n^{-1} = p(a)$ and $k|(n \times n^{-1} = p(a))$ then this implies $k|1$, but this a contradiction, hence $gcd(n, p) = 1$. Now we need to show that $\exists n^{-1}$ such that $n \times n^{-1} = p(a)$. Since $gcd(n, p) = 1$ then we can say $n(m)+p(l) = 1$ for some $a, b \in \mathbb{Z}$. Thus, $n(m) = 1-p(l) \Rightarrow n(m) = 1 mod p$. Let $n^{-1} = m mod p$ then this implies $n \times n^{-1} = 1 mod p$.

- Cyclic groups. A group is cyclic if there exists $a \in G$ such that for any $x \in G$ there is an integer $k \geq 1$ such that $a^k = x$. Basically if you raise one element in the group to any power you get a different (or the same) element which is in the same group.
- Semi-group only requires closure and associativity.
- Monoid only requires closure, associativity and an identity. Mo (NO Inverse) d.
- Isomorphism. Two groups $(G_1, \cdot)$ and $(G_2 . \star)$ are isomorphic if $\forall x, y \in G_1, f(x \cdot y) = f(x) \star f(y)$. Properties: If e is the identity of G_1 then $f(e)$ is the identity of G_2. If x^{-1} is the inverse of x in G_1 then $f(x^{-1})$ is the inverse of f(x) in G_2.
- Algebraic Structures $(A, \oplus, \bullet)$. $\bullet$ distributes over $\oplus$ if $\forall x, y, z \in A, x \bullet (y \oplus z) = (x \bullet y) \oplus (x \bullet y)$
- Rings. A ring is an algebraic structure that satisfies the following conditions:
  - $(A, \oplus)$ is an abelian group
  - $(A, \bullet)$ is a semi-group.
  - Operation $\bullet$ distributes over $\oplus$.
  - (Commutative rings have the operation $\bullet$ being commutative as well)
  - (A ring with unity has an identity for operator $\bullet$)
  - (A division Ring is a ring with unity and every element has an inverse, such that $x \bullet x^{-1} = 1$)
- Integral Domain is an algebraic structure that is a commutative ring with unity and satisfies the following property: $\forall x, y \in A, x \bullet y = 0 \Rightarrow x = 0$ or $y = 0$
- Field (Commutative division Ring) is an algebraic structure that follows the following properties:
  - $(A, \oplus)$ is an Abelian group
  - $(A - (0)), \bullet)$ is an Abelian group
  - $\bullet$ distributes over $\oplus$

GRAPH THEORY

- Definitions
  - Graph - An ordered pair $G = (E, V)$, where V is a set of vertices and E is a set of edges. E is a two-element subset of V.
  - Simple graph - A graph without loops and multi-edges

  - Path/Walk - Alternating sequence of links and nodes
  - Trial - A walk without repeated edges. (basically equal to an Eulerian path.
  - cycle - first vertex is the same as the last vertex
  - Eulerian Path - contains every link in the path exactly once
    * E.P exists if there are no more than two nodes of odd degree. If this is the case, it starts at one of the odd degree nodes and ends at the other.
    * If all nodes are of even degree, the paths are cycles.
    * If there are two odd nodes it might be referred to as a semi-Eulerian path.
  - Hamiltonian Path - contains every node exactly once
  - degree - number of links anchored to a node
  - Degree Sequence - Non-increasing sequence of node degrees
  - Undirected graph - edges have no orientation
  - Directed graphs - edges have a direction
  - weighted graph - edges contain a weighting, i.e. a number that describes their cost, distance or whatever is being analyzed.
  - Planar graph - none of the edges intersect on a 2D plane.
  - face - area enclosed by edges.
  - tree - connected graph without cycles
  - Complete Graph - every vertex is connected to every other vertex by an edge
  - clique - fully connected (complete) sub-graph of an undirected graph.
  - Adjacency matrix of a graph - A matrix representing all the adjacent vertices in a graph.

- Ore's theorem. A theorem that gives a sufficient condition for a graph to be Hamiltonian - $deg(v) + deg(w) \geq n$ for $n \geq 3$
- Efficiency of algorithms
  - Prim's is quicker than Kruskal's algorithm, because Kruskal requires the weightings to be sorted from smallest to largest first.
- Kruskal's algorithm
  - List weightings from smallest to largest in a list
  - Pick smallest weighting, draw its edge (as long as it does not form a cycle)
  - Remove from list and repeat steps 1 to 3.
- Prim's algorithm - Pick a node.

  - Create an empty set of visited nodes
  - Start at an arbitrary node and add it to the list, find the next nearest node and add it to the list.
  - Find the next nearest node from the nodes in the list.
  - Repeat process until every node is in the list.

- $\underline{K_n}$ is a complete graph with n nodes.