

Knowledge:

Knowledge enables us to act properly according to situation. Knowledge is necessary to show intelligent behavior. To show intelligent behavior, machine must have some way to represent the knowledge. To express the knowledge, we must have some language. We must empower the machine with the knowledge of particular domain in which machine is expected to act. The language must be understood by machine, and there must be some method to use it.

Inference mechanism reads the environment, use the knowledge and show certain action. Any language must have some grammar or rules so that it can be understood or interpreted in particular way. This helps the inference machine to understand it. It must have some syntax and semantics. Pen cuts road is syntactically correct but semantically incorrect. Inference is important to use knowledge and derive new facts from existing.

Logic:

Logics are formal languages for representing information such that conclusions can be drawn: It has syntax and semantics:

- Syntax defines rules or grammar required for correct use and interpretation of language. It specifies the symbols, operators, variables we can use in the language.

- Semantics define the meaning of sentences. That allows us to assign value to sentence interpreted in certain world

Propositional logic:

Proposition is simple statement or sentence which may be true or false and propositional logic is the language that allows to represents sentence, operate these facts and deduce new facts. Similar to any language, it has syntax and semantics.

Syntax

The syntax or rules help us to represent statements, interpret it, make the knowledge base system, deduce new facts from the existing ones etc. It allows us to represent or label the sentences in symbol. We can use set of sentences in propositional logic. The sentence may be atomic or compound sentence. For eg. Ram is tall. This is a proposition and propositional logic allows us to label this statement symbolically by P.

Atom: Single proposition is atom. For eg. Today is Sunday. Constants like TRUE or FALSE are also taken as atoms.

Literal: A single proposition or its negation is called literal.

Compound sentence: Two or more sentence combined through various connectors are called compound statements. Proposition logic has various logical operators or connectors like V (or) \wedge (and) \neg (Not) \rightarrow (implication). Its value also may be true or false based on interpretation in certain world.

Let P and Q denotes some proposition, then $P \vee Q$ is also a sentence whose value is be true if one of P or Q is true otherwise false.

P	Q	$P \vee Q$
T	T	T
T	F	T
F	T	T
F	F	F

$P \wedge Q$ is also a sentence whose value is be true when both of P and Q are true otherwise false.

P	Q	$P \wedge Q$
T	T	T
T	F	F
F	T	F
F	F	F

$\neg P$ is the sentence whose value will be true if P is false otherwise true.

P	$\neg P$
T	F
F	T

$P \rightarrow Q$ is a sentence that states the if P is true then Q is true. It can be written as $\neg P \vee Q$ as well as $\neg Q \rightarrow \neg P$.

P	Q	$\neg P$	$\neg P \vee Q$	$P \rightarrow Q$
T	T	F	T	T
T	F	F	F	F
F	T	T	T	T
F	F	T	T	T

$P \leftrightarrow Q$ is the sentence states that P is true if and only if Q. It is similar to P XNOR Q i.e. it is ANDing of $P \rightarrow Q$ and $Q \rightarrow P$

P	Q	$Q \rightarrow P$	$P \rightarrow Q$	$P \leftrightarrow Q = (P \rightarrow Q) \wedge (Q \rightarrow P)$	P XNOR Q
T	T	T	T	T	T
T	F	F	F	F	F
F	T	T	F	F	F
F	F	T	T	T	T

Semantics:

Interpretation of proposition in certain world provides the semantic or meaning of statement and assigns value to the statement that may be true or false. A statement that is true in some world may be false in some other world. It means interpretation is mapping of the given proposition in

particular world to check its meaning or semantics. Interpretation is also defined as the assignment of Boolean constant to Boolean variable or proposition.

Satisfiability:

If a sentence is evaluates to true under at least one interpretation, then the sentence is said to have satisfiable. To check the satisfiability, we assign Boolean values to propositions and see whether it is true at least for once.

Validity:

A proposition is said to be valid when it is always true irrespective of the world of its interpretation. For eg. $P \vee \neg P$ is always true thus it is a valid sentence or proposition such a statement is also called tautology.

P	Q	$\neg P$	$P \vee Q$	$P \vee \neg P$
T	T	F	T	T
T	F	F	T	T
F	T	T	T	T
F	F	T	F	T

Here, $P \vee Q$ is satisfiable proposition and $P \vee \neg P$ is a valid proposition.

Entailment:

If set of sentences S is true under all interpretation and a sentence S1 is also true under the same interpretations then, set of sentences logically entails the sentence S1. It means if we can deduce a sentence or fact S1 from given set of sentences or facts and rules S then the sentence S1 follows the set of sentences S under some interpretation.

Inference rules allows you to deduce the new things from existing ones. We have two popular inference rules namely modus ponens and resolution.

Deductive reasoning:

Derives conclusions or new facts by combining given premises(facts) and rules.

Inductive reasoning:

Derives new rule by using set of premises, rules and conclusions.

Abductive reasoning:

Derives premises from set of rules, and conclusions.

Modus ponens:

This rule say that if $P \rightarrow Q$ and P is given (true) then we can infer or conclude that Q is true. This rule is written as follows. $\frac{P \rightarrow Q, P}{Q}$, Where upper one is premises and down one is conclusion.

Premises means given facts or conditions.

Horn Clause:

$P_1 \wedge P_2 \wedge P_3 \dots \wedge P_k \rightarrow Q$ in which no ORing is possible in LHS and only atomic sentence in RHS. No all proposition can be converted into Horn Clause. Horn clause is used as a rule for forward and backward chaining.

Forward Chaining:

Knowledge Base consists set of facts and set of rules. Facts are the proposition in knowledge base and rules are in the form of implication. Facts in knowledge base are always true. In forward chaining, we start facts and rules and deduce new facts. The new fact is kept in knowledge base and again new facts are deduced. This process is continued till we get our goal. Suppose we have following facts in knowledge base.

F1: A

F2: B

F3: D

Also suppose we have following rules.

R1: $A \wedge B \rightarrow C$

R2: $C \wedge D \rightarrow E$

R2: $C \wedge F \rightarrow G$

Now deduce if E is a fact.

From fact F1 and F2 and Rule R1, C is true, thus kept in knowledge base.

From facts C, D and rule R2, E is fact. Hence proved.

Backward Chaining:

In backward Chaining, We start from goal, and check whether the RHS of rules in knowledge base contains the goal propositions. If so, we replace the goal proposition with LHS of the rule. This is done till we get the facts in our goal. This is done in most of logic programming language. Backward chaining is more powerful as it only focus on the goal. It does not compute any unnecessary goals. For eg.

Suppose we have following facts in knowledge base.

F1: A

F2: B

F3: D

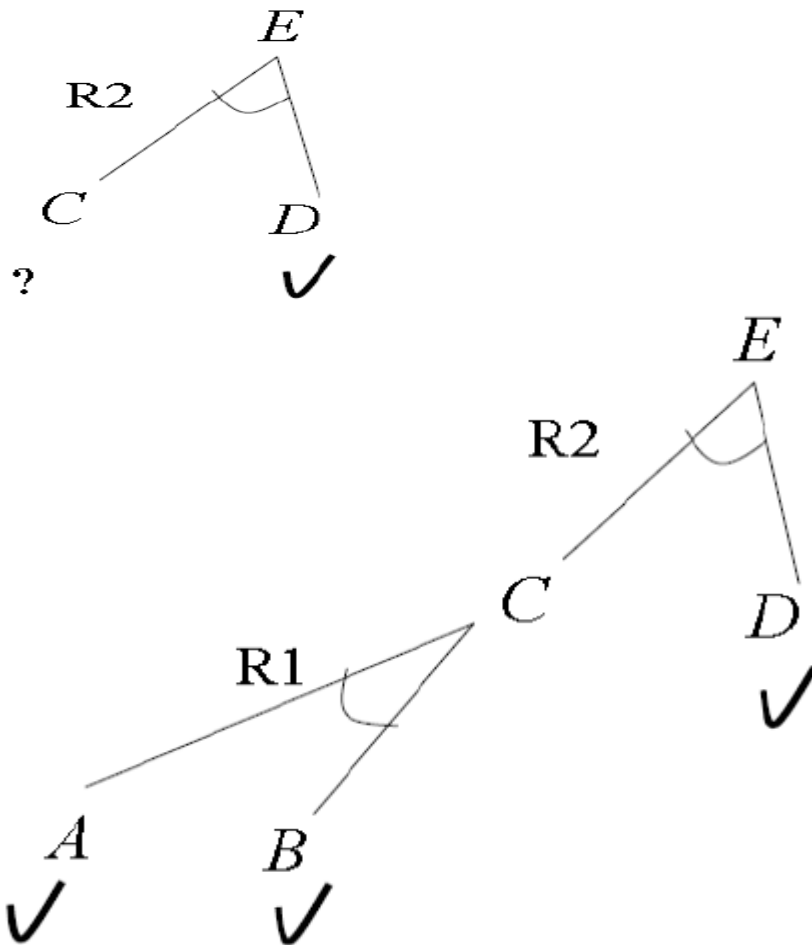
Also suppose we have following rules.

R1: $A \wedge B \rightarrow C$

R2: $C \wedge D \rightarrow E$

R2: $C \wedge F \rightarrow G$

Now deduce if E is a fact using backward chaining.



Conjunctive Normal form: It has compound terms with disjunction operators which are ultimately combined through AND operator is called conjunction. For eg. $(P \vee T) \wedge Q \wedge (\neg N \vee M)$

Disjunctive Normal form: The combination of compound sentences that may contain conjunctive operator through OR operator is called Disjunctive Normal form.

For eg. $(P \wedge S) \vee (Q \wedge R)$.

Clause: The combination of literals through OR operator is called clause. $P \vee Q, P, \neg C$

To convert $\neg (A \rightarrow B) \vee (C \rightarrow A)$ compound proposition into clause, following steps must be taken.

1. Eliminate the implication $\neg (\neg A \vee B) \vee (\neg C \vee A)$
2. Eliminate the Double negation sign $(A \wedge \neg B) \vee (\neg C \vee A)$
3. Apply distributive or associative laws to get conjunctive normal forms. $(A \vee \neg C \vee A) \wedge (\neg B \vee \neg C \vee A) ((A \vee \neg C) \wedge (\neg B \vee \neg C \vee A))$
4. Get the set of clauses.
 $(A \vee \neg C)$
 $(\neg B \vee \neg C \vee A)$

Clausal form allows us to prove the fact given the set of facts through the process of resolution. In fact process of proving a fact with given sets of fact is called entailment.

Convert followings in CNF:

1. $(P \wedge Q) \vee (P \wedge \neg Q)$
2. $\neg (\neg P \wedge (Q \vee \neg (R \wedge S)))$
3. $\neg (Q \wedge R) \vee S$
4. $B \leftrightarrow (A \vee C)$

Represent followings into propositional logic.

1. Ravi works hard. If ravi works hard, then he is a dull boy. If he is a dull boy, he will not get a job. Prove that Ravi won't get job.
2. If it neither rained nor foggy, the sailing race will be held and lifesaving demonstration will go on. If sailing race will go on, the trophy will be awarded. Trophy was not awarded. Prove that, it rained

Resolution:

Resolution is another method of inference. This rule says that if $P \vee Q$ is true and $\neg P$ is given (true), then Q is true. This rule is written as $\frac{P \vee Q, \neg P}{Q}$. Resolution can be also written as $\frac{P \vee Q, \neg Q \vee R}{P \vee R}$

$$\frac{\neg P \rightarrow Q, Q \rightarrow R}{\neg P \rightarrow R}$$

Steps in resolution.

- Convert the given proposition into clausal form.
- Negate the sentence to be proved.
- Combine the clauses into set.
- Iteratively apply the resolution to the clausal sets.
- If we got the null clause then it lets us in the situation of contradiction and it proves the given fact.

With following KB prove that S can be inferred with resolution but not with modes ponens.

$P \rightarrow Q, Q \rightarrow S, \neg P \rightarrow R, R \rightarrow S$

Exercise:

1. Represent the following proposition in propositional logic labels and deduce some conclusions.

If the unicorn is mythical, then it is immortal but if it is not mythical, then it is mortal mammal. If the unicorn is either immortal or a mammal, then it is horned. The unicorn is magical, if it is horned. Check if the unicorn is mythical, mammal or mortal.

Let

P : unicorn is mythical.

Q: unicorn is mortal.

R: unicorn is mammal.

S: unicorn is horned.

T: unicorn is magical.

From given proposition,

$P \rightarrow \neg Q, \neg P \rightarrow Q \wedge R, \neg Q \vee R \rightarrow S, S \rightarrow R$

2. Represent following English sentences into propositional logic.

It rains in July.

This book is not costly.

If it rains today and one does not carry umbrella, he will be drenched.

3. Represent following into propositional logic and deduce some conclusion.

Mammals drink milk.

Man is mortal.

Man is a mammal.

Tom is a man.

Prove Tom drink milk and tom is mortal.

4. Consider the knowledge base $(B \leftrightarrow (A \vee C)) \wedge \neg B$. Prove that $\neg A$ can be inferred from this KB by using resolution.

5. $(P \vee R) \wedge (R \rightarrow Q) \wedge (P \rightarrow T) \wedge \neg S \wedge (S \rightarrow T)$ prove that Q is true through resolution.

6. $(P \wedge Q), (P \rightarrow R), (Q \wedge R) \rightarrow S$ are in knowledge base, prove that S is true using resolution

7. $P \rightarrow Q$

$L \wedge M \rightarrow P$

$B \wedge L \rightarrow M$

$A \wedge P \rightarrow L$

$A \wedge B \rightarrow L$

A

B

Prove that Q can be inferred from above KB using forward and backward chaining.

8. Prove that $\text{smoke} \rightarrow \text{smoke}$ is valid and satisfiable but $\text{smoke} \rightarrow \text{fire}$ is satisfiable but not valid.

First Order Logic:

First order logic is more powerful language than propositional logic. In first order logic, we use predicate to represent the statements and predicate take one or more arguments. The predicate may be true or false. It also has syntax and semantics. This logic assumes that world contains object that have some function, properties and some relations. Thus first order logic is more powerful as it can express the object, its properties, its functions and relation to other object. First order logic can quantify the objects also.

Syntax:

We can use followings to represent statements in first order logic.

Constant: we can use constants like a, 5, lalitpur. In fact it represents the objects of real world.

Variables: Variables can be used in capital letters.

Predicate: we can use predicate that take arguments and return true or false value when instantiated with the variable of some domain. For eg. $\text{Mother}(x,y)$. The predicate can represent

some property of object or relationship between various objects. For eg tall(ram). Here ram is object and tall is its property. Mother(x,y) depicts relationship between x and y objects. Predicate can take function as its argument also.

Function: We can use function that take arguments and return some value which may numeric, non-numeric etc. function represents the function of an object.

Everyone loves their mother. $\forall x \exists y \text{Mother}(x, y) \text{loves}(x, y)$ Here loves is predicate and so is mother.

$\forall x \text{loves}(x, \text{Mother}(x))$ Here mother is function that returns mother of x and loves is predicate. In FOL, functions will never appear outside the predicate.

Different types of sentences like atomic sentence, complex sentence, quantified sentences.

Atomic sentence: It is the predicate that takes the arguments like constant, variable, or functions. For eg mother(x,y).

Complex sentence: It is combination of two or more atomic sentences through connectives like $\wedge, \neg, \vee, \rightarrow, \leftrightarrow$. For eg $\text{Mother}(x, y) \vee \text{loves}(x, y)$

Quantified sentences contains quantifiers like \forall, \exists . It is because of quantifiers which makes inferencing in FOL makes very difficult.

Not all students take history and biology.

Student(x) : x is a student

Takes(x,y) : subject x is taken by student y.

$$\forall x \text{student}(x) \rightarrow \text{Takes}(\text{History}, x) \wedge \text{Takes}(\text{Biology}, x)$$

Express following in first order logic.

Cats are mammals. $\forall x \text{cat}(x) \rightarrow \text{mammal}(x)$

Jane is a tall surveyor. $\text{Tall}(\text{jane}) \wedge \text{surveyor}(\text{jane})$

A nephew is a sibling's son. $\forall xy (\text{nephew}(x, y) \rightarrow \exists z (\text{sibling}(x, z) \wedge \text{son}(y, z)))$

Everybody loves somebody. $\forall x \exists y \text{loves}(x, y)$

Nobody loves Jane. $\forall x \neg \text{loves}(x, \text{Jane}) \rightarrow \neg \exists x \text{loves}(x, \text{Jane})$

Everybody has a father.

All men are mortal. $\forall x \text{man}(x) \rightarrow \text{mortal}(x)$

Some birds cannot fly. $\exists x \text{bird}(x) \wedge \neg \text{fly}(x)$

At least one planet has life on it. $\exists x \text{planet}(x) \wedge \neg \neg \text{life}(x)$.

If ram is man, then ram is mortal. $\text{Man}(\text{ram}) \rightarrow \text{mortal}(\text{ram})$

All dogs are faithful. $\forall x \text{dog}(x) \rightarrow \text{faithful}(x)$

Some dogs bark. $\exists x (\text{dog}(x) \wedge \text{bark}(x))$

All dogs have four legs. $\forall x \text{dog}(x) \rightarrow \text{fourleg}(x) \quad \forall x \text{dog}(x) \rightarrow \text{leg}(x, 4)$

All barking dogs are irritating. $\forall x \text{dog}(x) \wedge \text{bark}(x) \rightarrow \text{irritating}(x)$

No dogs purr. $\neg \exists x (\text{dog}(x) \wedge \text{purr}(x)) \quad \text{dog}(x) \rightarrow \text{faithful}(x)$
 $\forall x \neg$

Fathers are male parent with children. $\forall x \text{father}(x) \rightarrow \text{male}(x) \wedge \text{haschildren}(x)$
Students are people who are enrolled in courses.

Inference rules: It tells what can be deduced from what.

Universal Elimination: $\forall x \text{likes}(x, \text{ice-cream})$ when x is substituted with ram , we get $\text{likes}(\text{ram}, \text{iceram})$ which is proposition. Universal elimination is done by using ground term i.e. some constant which contains no variables and will not have any further instantiation..

Existential Elimination:

$\exists x \text{likes}(x, \text{icecream})$ we can eliminate existential elimination with skolem constant which will not appear in knowledge base elsewhere. For eg. $\text{Likes}(\text{ram}, \text{icecream})$ is elimination of existential quantifier. That does not appear in knowledge base. Skolem constant is also a ground term as it will have no further instantiation. This process is also called skolemization.

Existential Introduction:

We can convert $\text{likes}(\text{monalisa}, \text{ice-cream})$ $\exists x \text{likes}(x, \text{ice-cream})$ as we know that monalisa likes ice-cream which means there exists some x who likes ice-cream.

Horn sentences:

Any atomic sentences are horn sentences. For eg $\text{likes}(x, \text{flower})$ is a horn sentence. Implication with conjunction of predicates at LHS with no existential quantifier and atomic sentence at RHS is called horn sentence without any existential quantifier. For eg. $\forall x, y \text{persquare}(x) \wedge \text{prime}(x) \rightarrow \text{square}(x)$

Unification:

It is a process of finding a particular substitution that makes two atomic sentences identical. Two atomic sentences like $\text{prime}(x)$ and $\text{prime}(7)$ can be identical when x is substituted with 7. It means when these two sentences are unified, the x is replaced with 7. It is formally written as $\text{UNIFY}(\text{prime}(x), \text{prime}(7)) = \{x/7\}$. Unification of $\text{prime}(x, 7)$ and $\text{prime}(47, x)$ is impossible. The unification is necessary when we use modus ponens for reasoning. It is also required for resolution process. Unification provides a special fact for some particular subject. The unification must be unrestricted. Unification must be as our requirement and we must use most general unifier.

Substitution:

It replaces the variables with constant. It is used during unification of two predicates. For eg. $\text{SUBST}(\{x/49\}, \text{perfect}(x)) = \text{perfect}(49)$.

AND elimination:

$\text{Prime}(x) \wedge \text{composite}(x)$ can be written as $\text{prime}(x), \text{composite}(x)$ by eliminating AND operator. AND operator can be introduce if two predicates are true.
Resolution in FOL

Duality

All men are mortal can always be written as no man is immortal. $\forall x \text{man}(x) \rightarrow \text{mortal}(x) = \neg \exists x (\text{mortal}(x) \wedge \neg \text{mortal}(x))$

There exists birds that can fly. It is not the case that all birds cannot fly. $\exists x (\text{bird}(x) \wedge \text{fly}(x)) = \neg \forall x (\text{bird}(x) \rightarrow \neg \text{fly}(x))$

Represent following statements in first order logic and check if Traitorix is a criminal using forward and backward chaining.

Law says that it is a crime to sell potion formula to hostile nations. The country Rome, an enemy of Gaul has acquired some potion formulas and all of its potion formulas was sold by Druid Traitorix. Traitorix is a Gaul.

Gaul(x)

Hostile(z)

Potion(y)

Criminal(x)

Sells(x,y,z).

Owns(x,y)

$\text{gaul}(x) \wedge \text{potion}(y) \wedge \text{sells}(x, y, z) \wedge \text{hostile}(z) \rightarrow \text{criminal}(x)$

Hostile(Rome)

$\exists y (\text{potion}(y) \wedge \exists (y, \text{rome}))$

$\text{potion}(y) \wedge \text{owns}(y, \text{rome}) \rightarrow \text{sells}(\text{Traitorix}, y, \text{rome})$

$\forall x$

Gaul(Traitorix)a

Forward chaining

$\exists y (\text{potion}(y) \wedge \exists (y, \text{rome}))$

$\text{potion}(p) \wedge \exists (p, \text{rome})$ (Eliminating existential quantifier with p that is not in knowledge).

$\forall x (\text{potion}(y) \wedge \text{owns}(y, \text{rome})) \rightarrow \text{sells}(\text{Traitorix}, y, \text{rome})$

$\text{potion}(p) \wedge \text{owns}(p, \text{rome}) \rightarrow \text{sells}(\text{Traitorix}, p, \text{rome})$ (Unification by substituting y with p)

$\text{gaul}(x) \wedge \text{potion}(y) \wedge \text{sells}(x, y, z) \wedge \text{hostile}(z) \rightarrow \text{criminal}(x)$

$\text{gaul}(\text{Traitorix}) \wedge \text{potion}(p) \wedge \text{sells}(\text{Traitorix}, p, \text{rome}) \wedge \text{hostile}(\text{rome}) \rightarrow \text{criminal}(\text{Traitorix})$

....b

(Unification by substituting {x/Traitorix}, {y/p}, {z/Rome})

Hostile(Rome) (substitution {z/Rome}) c

Sells(Traitorix,p,Rome) d

Potion(p)e

From a,b,c,d,e

Criminal(Traitorix)

Modus ponens is not a complete inferring technique as it can't deduce the negation of some fact. And can't deduce new facts. It means if any implication in knowledge base has negation then it can't be converted to horn clause and modus ponens can't be used. It means it can't deduce the fact which actually can be deduced. For eg.

$\forall x p(x) \rightarrow Q(x)$
 $\forall x Q(x) \rightarrow S(x)$
 $\forall x \neg p(x) \rightarrow R(x)$
 $\forall x R(x) \rightarrow S(x)$

From above knowledge base, we must be able to deduce $S(A)$ is true. But using modus ponens, we can't.

Steps to convert into Conjunctive normal form.

- Eliminate implication.
- Move negation inwards
- Push the quantifiers left as far as possible.
- Eliminate the existential quantifier through skolemization.
- Drop all universal quantifiers.
- Distribute \wedge over \vee

Steps to deduce new facts through resolution

- Negate the sentence to be proved and make it clause.
- Combine all the clauses into a set.
- Iteratively apply the resolution to the set and add the resolvent to the set.
- Continue till the null clause is obtained.

For eg. Can we infer someone smiling from following sentences?

All people who are graduating are happy.

All happy people smile.

Someone is graduating.

$\forall x \text{graduating}(x) \rightarrow \text{happy}(x)$ 1

$\forall x \text{happy}(x) \rightarrow \text{smiling}(x)$ 2

$\exists x \text{graduating}(x)$ 3

Goal is to prove

$\exists x \text{smiling}(x)$

Negate the goal

$\neg \exists x \text{smiling}(x)$

Converting 1 and 2 into clausal form we get

$\forall x \neg \text{graduating}(x) \vee \text{happy}(x)$ 4

$\forall x \neg \text{happy}(x) \vee \text{smiling}(x)$ 5

$\exists x \text{graduating}(x)$ 3

$\neg \exists x \text{smiling}(x)$ 6

4,5,6,7 are clauses we have.

Reduce the scope of negation in 6

$\forall x \neg \text{smiling}(x)$ 7

Standardize the variables apart.

Move all quantifiers to the left as much as possible. This must be done when there are more than 1 quantifier and not at the left of clause. Here is no change at all as all quantifiers are at the leftmost.

Eliminate the existential quantifier using the skolem constant. Here ram is a skolem constant.

$\text{graduating}(\text{ram})$

Drop all the universal quantifiers.

$\neg \text{graduating}(x) \vee \text{happy}(x)$ a

$\neg \text{happy}(x) \vee \text{smiling}(x)$ b

$\text{graduating}(\text{ram})$ c

$\neg \text{smiling}(x)$ d

Resolving b and d we get

$\neg \text{happy}(x)$

This is resolved with a to get

$\neg \text{graduating}(x)$

This is resolved with c to get null clause. This proves that our supposition is wrong i.e. smiling(x) is true.

Reasoning under uncertainty:

Suppose we have following rule in our knowledge base.

$\forall \text{symptom}(p, \text{toothache}) \rightarrow \text{disease}(p, \text{cavity})$ This rule is not correct because toothache can be caused by many causes. Toothache can be caused through gum disease, hit by somebody. These all causes must be specified comprehensively in the rule. Some causes may not be known even. Thus specifying all these causes is a problem. In such a case which is uncertain about the causes, we can use probability. Probability is very powerful tool for reasoning in uncertainty. We compute the probability of some event and infer something from this probability.

The reasons for using probability reasoning under uncertainty are as follows

Specification becomes too large.

All antecedents are not known.

Some antecedents are not measurable.

Probability refers if a person is fat or not as we don't know the person. **Fuzzy logic** refers how fat is the person. Fat is not a true, false value in fuzzy logic. It is gradation. Possibility of being fat is fuzzy.

Basics of probability

$$0 \leq P(A) \leq 1$$

$$P(\text{true}) = 1$$

$$P(\text{false}) = 0$$

Union probability of two independent events A and B is $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$.

Mutually exclusive events can't occur simultaneously.

Prior probability (unconditional probability):

It corresponds to the degree of belief in absence of any other belief.

For eg. $P(\text{cavity}) = 0.1$

Conditional Probability:

Occurring of some events may depend upon the occurring of other events. In such case we must calculate the conditional probability. Conditional probability of an event is defined as the it's probability given other events. Conditional probability is used for dependent events.

For eg. $P(\text{cavity}|\text{toothache}) = 0.8$ means probability of cavity when there is a toothache is 0.8.

Independence among events:

Some events which does not affect the event of which probability is being calculated, then such events can be dropped. For eg $P(\text{john calls}|\text{alarm, wakeup})$ here wakeup event does not affect the probability of john calls thus wakeup term can be dropped.

For eg if A is independent of B then

$$P(A|B)=P(A).$$

$$P(B|A)=P(B).$$

$$P(A \wedge B)=P(A)*P(B).$$

Bayes rules:

Bayes rule is very important for computing conditional probability for dependent events and reasoning under uncertainty. Bayes rules states that

$$P(A|B)=P(B|A)*P(A)/P(B)$$

This can be proved like this

$$P(A \wedge B)=P(A|B)*P(B) \dots\dots\dots 1$$

$$P(B \wedge A)=P(B|A)*P(A) \dots\dots\dots 2$$

From 1 and 2,

$$P(A|B)*P(B)=P(B|A)*P(A)$$

$$\text{Or } P(A|B)=P(B|A)*P(A)/P(B)$$

Joint probability is the method of representing the probability of all events. Joint probability distribution encompasses all the probabilities of various combinations of different variables. It also allows us to compute the probability of some events provided that other events are true or false. In fact advantage of **Full joint probability distribution** is it enables us to answer all types of queries. Joint probability of two independent events A and B is $P(A \wedge B)=P(A)*P(B)$.

$$\begin{aligned} P(X_1, X_2, \dots, X_n) &= P(X_n | X_{n-1}, X_{n-2}, \dots, X_1) * P(X_{n-1}, \dots, X_1) \\ &= P(X_n | X_{n-1}, X_{n-2}, \dots, X_1) * P(X_{n-1} | X_{n-2}, \dots, X_1) P(X_{n-2}, \dots, X_1) \\ &= \prod_{i=1}^n P(X_i | X_{i-1} \dots X_1) \\ &= P(X_i | \text{parents}(X_i)) \end{aligned}$$

Consider following example.

	toothache		\neg toothache	
	catch	\neg catch	catch	\neg catch
Cavity	0.108	0.012	0.072	0.008
\neg cavity	0.016	0.064	0.144	0.576

$$P(\text{cavity})=0.108+0.012+0.072+0.008=0.200$$

$$P(\text{toothache} \vee \text{cavity})= 0.108+0.012+0.072+0.008+0.016+0.064=0.280$$

Start with the joint distribution:

	<i>toothache</i>		\neg <i>toothache</i>	
	<i>catch</i>	\neg <i>catch</i>	<i>catch</i>	\neg <i>catch</i>
<i>cavity</i>	.108	.012	.072	.008
\neg <i>cavity</i>	.016	.064	.144	.576

Can also compute conditional probabilities:

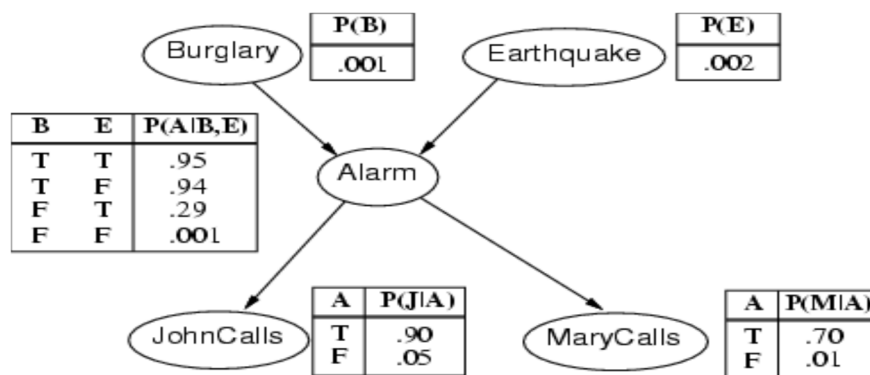
$$\begin{aligned}
 P(\neg \text{cavity} | \text{toothache}) &= \frac{P(\neg \text{cavity} \wedge \text{toothache})}{P(\text{toothache})} \\
 &= \frac{0.016 + 0.064}{0.108 + 0.012 + 0.016 + 0.064} = 0.4
 \end{aligned}$$

Belief Network/Bayesian Network

Belief network is the attempt to understand the cause effect relationship with their probability. It means it is representation of causal connections between different events with associated conditional probability. Belief network which is also known as bayesian network or causal network can represent the uncertain knowledge and we can reason under uncertainty through conditional probability of the events. There are two phases reasoning using belief network. First phase is to learn the belief network and second phase is use of belief network. The learning of belief network can be done with experimental data. This is called inductive reasoning. A belief network is a graph with followings

- Nodes: It is set of random variables.
- Directed links: The intuitive meaning of link from a node X to another node Y is that X has direct influence on Y.
- The graph has no directed cycles (DAG).
- Each node has conditional probability that quantifies the effect of its parent node.

For eg Burglary alarm at home. Burglary alarm can be caused by somebody at the door, short circuit, wind, animals etc. The cause may be infinite and listing of them in rule is not possible. Such a rule that does not contain all causes can't be tautology and can't be useful. The alarm is reliable at detecting burglary. It responds at minor earthquake also. Two neighbours john and mary can call police hearing the alarm. John always calls the police when he hears the alarm. He even calls the police when he is confused with the telephone rings with alarm. But mary likes loud music and sometimes misses the alarm. Belief network of this scenario is given below.



$$P(J \wedge M \wedge A \wedge \neg B \wedge \neg E) = P(J|A) * P(M|A) * P(A|\neg B, \neg E) * P(\neg B) * P(\neg E) = .90 * .70 * .001 * (1 - .001) * (1 - .002)$$

Construction of Bayesian Network.

1. Choose the set of relevant variables X that describes the domain.
2. Choose an ordering for the variables (very important step).
3. While there are variables left,
 - a. Pick a variable X and add a node for it.
 - b. Set $\text{parent}(X)$ to some minimal set of existing nodes such that the conditional independence property is satisfied.
 - c. Define the conditional probability table for X .

Ordering must be in such a way that causal effect relationship is in proper direction. If ordering is correct then we can learn the conditional probability table efficiently. Improper ordering increases size of probability table. The proper ordering of variables maximizes the use of conditional independence. Improper ordering results very confusing conditional probability that does not exist in practice. Probability table may be known or extracted from experimental data.

More on conditional independence.

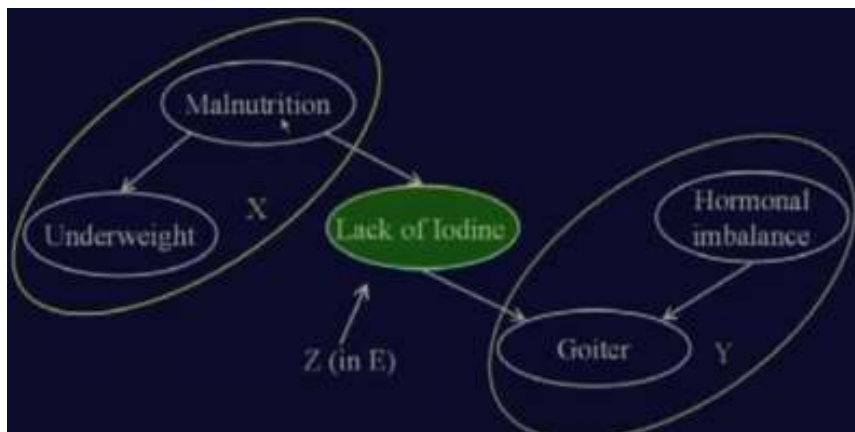
If information about parent node is not known then a node is not conditionally independent of its child, siblings and ancestors.

A set of nodes E d-separates two set of nodes X and Y if every undirected path from a node in X to a node in Y is blocked given E .

A path is blocked given a set of nodes E if there exists a node Z in the path for which one of the following holds.

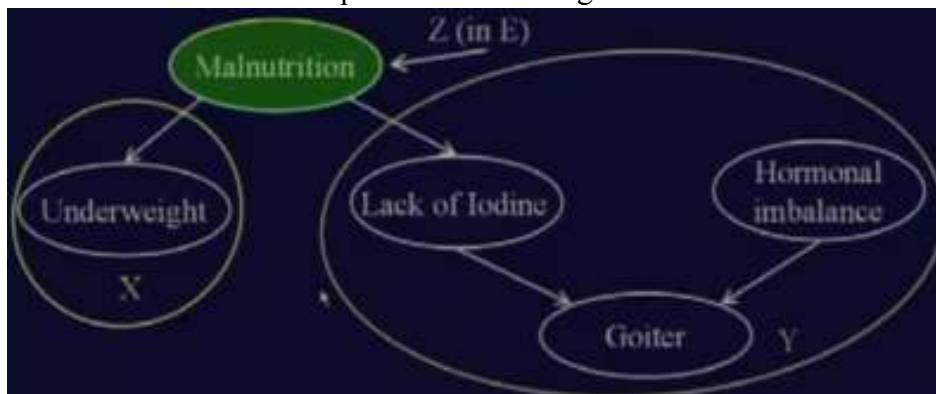
- Z is in E and Z has one arrow on the path leading in and one leading out.

Here the evidence lack of iodine has one path leading in and one path leading out of it. Now this node d-separates other nodes and other nodes becomes independent of each other. Now goiter only depends upon hormonal imbalance and lack of iodine. It means we can infer goiter or compute its conditional probability given hormonal imbalance and lack of iodine.



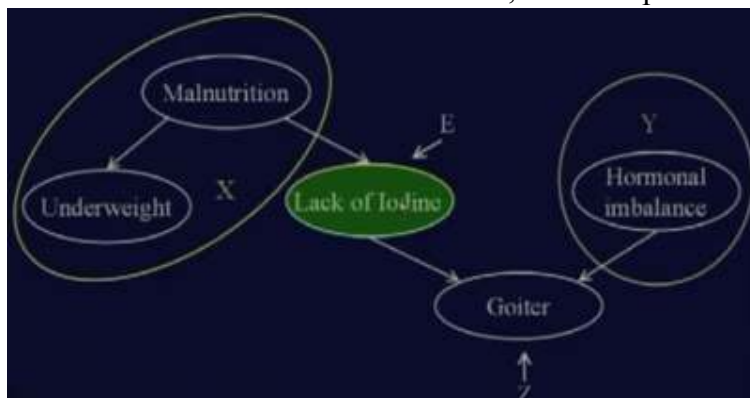
If every undirected path from the set of nodes in X and set of nodes in Y is d-separated by a given set of evidence nodes E, then X and Y are conditionally independent given E. This allows us to restrict our computation to subset of belief network.

- Z is in E and Z has both path arrows leading out.



Here malnutrition is evidence that has both paths leading out of it. Thus this also d-separates other nodes.

- Neither Z nor its descendants are in E, and both path arrows lead in to Z.



Here lack of iodine is evidence and goiter has both path leading in. Thus goiter d-separates other nodes.

Inferences in Belief Network.

After representing the uncertain knowledge in belief network, we can inference in belief network by making queries like what is probability of some event given some events. For this we compute the probability and infer the probability of some events. Following types are the inferences we can do in belief network.

Diagnostics Reasoning: from effects to causes i.e. abductive reasoning. It means it is computing the probability of parent given its child for eg. Given that john calls what is the probability of burglary?

$$P(B|JC)=0.016$$

Causal Inferences: From causes to effects that is deductive reasoning. It means it is computing the probability of a child given the parent. For eg. Given the burglary, what is the probability that john calls, mary calls.

$$P(JC|B)=0.86$$

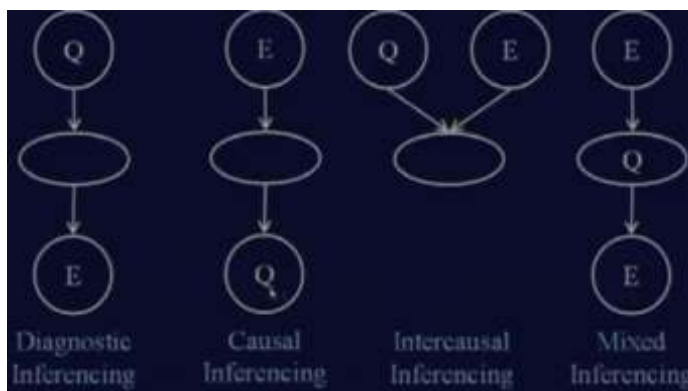
$$P(MC|B)=0.67$$

Intercausal inferences: It is the reasoning between causes of common events. It computes the probability of a parent given its siblings. Given alarm and earthquake what is the probability of burglary.

$$P(B|A,E)=0.003$$

Mixed inferences: Some causes and some effects are known. It means computing the probability of an event when its parent and child is known. For eg. Given john calls and no earthquake, what is probability of Alarm?

$$P(A|JC \wedge \neg E)=0.003$$



Default Reasoning:

It is paradigm where we make some default assumption about some facts. For eg. If we can see two wheel of a car from one side, then we can infer that there are two wheels in other side. Thus we make such assumption and add them in knowledge base. It is non-monotonic reasoning as the facts derived may change. Default rule can't be applied any time. For eg if we know that one tyre is stolen, then we can't say by default that there are 2 wheels at other side.

Rule based reasoning:

Rule based reasoning has some rules and facts based upon which new rules and facts are deduced. Thus modus ponens is a rule based reasoning. If precondition of a rule is matched then post condition of the rule is always true.