UNIVERSITY KASDI MERBAH OUARGLA
Faculty of New Information Technologies and Communication
Department of computer science and Information Technology

ACADEMIC MASTER Thesis

Domain : Computer science and Information Technology
Track : Computer science
Specialty : Fundamental ,industrial

Topic :

Semantic Network based appraoche for automatic generalization problem

Directed by : ALchaima SIKEBIR , Rabiha ABIDE

Supported publicly

the : 13/10/2020

Before the jury:

Pr. Mohamed lamine KHERFI	Professor	President	UKM Ouargl
Dr. Meriem KORICHI	M.C.B	Supervisor	UKM Ouargla
Dr. Khadra BOUANANE	M.C.B	Co-Supervisor	UKM Ouargla
Mr. Fouade BEKKARI	M.C.B	Examiner	UKM Ourgla

## *Acknowledgements*

First of all we would like to thank ALLAH the Beneficent, the Merciful. Praise be to ALLAH, Lord of the Worlds. ALLAH who guides us along the way.

Special thanks go to our dear parents and family for their support with all what They could give us.

As well, we would like to express our deepest gratitude to our supervisor Ms. Meriem KORICHI and Mrs. Khadra BOUANANE for providing valuable comments, wisdom, encouragements, guidance, giving us

The possibility to conduct this thesis with the maximum freedom and for her trust.

Many thanks to the jury members for accepting to evaluate our work.

Also, we appreciate the great aid and support from all the teachers of computer science department, Ouargla University.

Finally, we are indebted to thank whomever supported or helped us during our work.

Dedication

I dedicate this work to the source of my inspiration my dear parents

, To my dear sisters and brothers each by his name, to my dear supervisor, my friends and whomever I love or love me.

Chaima

Dedication

Firstly, I dedicate this work to my parents, who were and still are my support.

This work also I dedicate to every member of my family, to my friends

And to everyone who supported me and believe in me.

Rabiha

Abstract

One of the crucial goals for research in artificial intelligence is to understand the nature of human learning and implement learning capabilities in machine. Machine learning is the field of artificial Intelligence that is concerned with developing computational learning theories and constructing learning systems. Human learning is a very complex process and it has different forms (concept learning, word learning, behavior learning, etc.), from which we are interested in concept learning. Most of concept learning researches that have been done in artificial intelligence has consisted of either: using largely analytic techniques to classify inputs or supplying programs with examples and sometimes counter examples of a specified concepts and these programs determine the definition of concepts. Our work falls under the second category.

Our purpose is to construct a concept learning algorithm that can learn effectively from limited number of examples using semantic networks. And surpass the limits encountered in the existent approaches.

Key words: Artificial intelligence, Machine learning, concept learning, semantic networks

Résumé

L'un des objectifs cruciaux de la recherche en intelligence artificielle est de comprendre la nature de l'apprentissage humain et de mettre en œuvre des capacités d'apprentissage dans la machine. L'apprentissage automatique est le domaine de l'intelligence artificielle qui consiste à développer des théories d'apprentissage informatique et à construire des systèmes d'apprentissage. L'apprentissage humain est un processus très complexe et il prend différentes formes (apprentissage de concept, apprentissage de mots, apprentissage de comportement, etc.), à partir desquels nous nous intéressons à l'apprentissage de concept. La plupart des recherches sur l'apprentissage de concepts qui ont été effectuées en intelligence artificielle ont consisté soit: à utiliser des techniques largement analytiques pour classer les intrants, soit à fournir aux programmes des exemples et parfois des contre-exemples de concepts spécifiés et ces programmes déterminent la définition des concepts. Notre travail relève de la deuxième catégorie.

Notre objectif est de construire un algorithme d'apprentissage de concepts qui peut apprendre efficacement à partir d'un nombre limité d'exemples utilisant des réseaux sémantiques. Et dépasser les limites rencontrées dans les approches existantes.

Mots clés: intelligence artificielle, apprentissage automatique, apprentissage de concepts, réseaux sémantiques

ملخص

أحد الأهداف الرئيسية للبحث في مجال الذكاء الاصطناعي هو فهم طبيعة التعلم البشري وتنفيذ قدرات التعلم في الآلة. التعلم الآلي هو أحد مجالات الذكاء الاصطناعي والذي يهتم بتطوير نظريات تعلم حسابية وبناء أنظمة تعلم. يعتبر التعلم البشري عملية معقدة للغاية ولها أشكال مختلفة (تعلم المفاهيم، تعلم الكلمات ، تعلم السلوك ،... إلخ) ، ينصب اهتمامنا حول تعلم المفاهيم. تتكون معظم أبحاث تعلم المفاهيم التي تم إجراؤها في الذكاء الاصطناعي إما من: استخدام تقنيات تحليلية إلى حد كبير لتصنيف المدخلات أو تزويد البرامج بأمثلة وأحيانًا أمثلة مضادة لمفاهيم محددة وهذه البرامج تعرف المفاهيم. عملنا يندرج تحت الفئة الثانية.

هدفنا هو بناء خوارزمية تعلم مفاهيم يمكنها التعلم بفعالية من عدد محدود من الأمثلة باستخدام الشبكات الدلالية. وتجاوز الحدود التي واجهتها الأعمال القائمة في هذا الاطار.

الكلمات المفتاحية: الذكاء الاصطناعي ، التعلم الآلي ، تعلم المفاهيم ، الشبكات الدلالي

Contents

List of Figures

List of tables

GENERAL INTRODUCTION

Context and problematic

The issue of creating machines capable of simulating human behavior has monopolized a wide range of computer science researches for decades and still one of the most elusive subjects. The existence of such machines in our life, would help in solving many critical real world problems, starting from medical domain (eg. Machines that help in the early detection of brain tumor) [19], to military (eg. Creating intelligent defense systems) [18], industry (eg. Machines that detect immediately any fault of manufacturing during the process) [20], and so on.

Despite the considerable development in artificial intelligence, human conceptual knowledge has eluded machine systems in many aspects, The ability to learn concepts from a few number of examples is one of the dilemmas of human cognition, In contrast many machine learning approaches are the most 'data-hungry ' .How does people learn from a small number of examples? , Is the question that challenged researchers in psychology and Artificial intelligence for a long time and is still opened for contemporary researches .

Cognitive scientists cast the problem of learning concepts from only few examples as a generalization problem, building up on this principal several approaches in artificial intelligence literature have been proposed, The Bayesian framework for generalization is the famous one which has achieved a large success in many domains (medicine, bioinformatics, speech recognition…etc.), where it could generalize successfully from just a few examples [2] as it was extended to be used effectively in many application domains like image retrieval [17], however this later is successful only when the stimuli are represented in an appropriate psychological space or equivalently using a psychologically valid hypothesis space as well it takes in consideration only one type of semantic relations (Is_a)between concepts but ignores other types of relations ,what would cause loss of many valuable information ,and it may affect the accuracy of results as it may misled the generalization in some cases(eg. The target concept is not in the taxonomy of the specified semantic relation), this problem was also encountered in other classic approaches to concept learning, what encouraged us to search for solution. Starting from the idea that one of the essential aspect of any concept learning algorithm is the choice of a good learning space, and that the much the learning space is closer to human knowledge structure the much the learning algorithm will perform closer to human, we found in concept net semantic network the best and the appropriate learning space for our concept learning algorithm, because of the rich knowledge it provides and its exhaustive covering of several and different types of relations between concepts such the ones exiting in human brain.

Objective

We aim to propose a new concept learning algorithm for the problem of generalization in automatic concept learning that will surpass the limits of the Bayesian framework. The proposed algorithm is based on semantic network. We took the advantage of the rich structure of concept net to provide an appropriate learning space that would complete the search strategy of our algorithm.

Organization of thesis

This thesis consists of three chapters, general introduction and general conclusion:

In the first chapter we will give an overview about the domain where our contribution take a place in .in the second chapter we will present semantic networks which is the tool that the whole work is based on. The third chapter includes two sections, in the first section we will explain the principal of our algorithm and in the second one we will make an experimental evaluation for the performance of our algorithm and finally we will discuss the obtained results.

Chapter
01

Concept Learning and generalization

1. Introduction

Learning is one of the most tremendous capacities of human cognition, A few months after our birth we begin to learn about everything that is going around us, learn how to talk and to interact with our environment, then how to walk and so on. Cognitive scientists have argued that one a crucial capability underling several kinds of learning is the ability to take into account a number of specific observations then to extract the essential common features that characterize them, what defined the term "generalization". Generalization is prominently notion in cognitive theories of learning, it has been partially solved for different problems of learning ranging from learning fragments of spoken English to behavior learning … etc. This chapter consists of two main titles, first, we will start with generalization, under the second title we will look at concept learning.

2. What is generalization?

Several definition has been attributed to the term generalization, we can cite the following:

- "*A generalization is a form of abstraction whereby common properties of specific instances are formulated as general concepts or claims*[1]". Generalizations posit the existence of a domain or set of elements, as well as one or more common characteristics shared by those elements (thus creating a conceptual model).

- It is the conclusion of the meaning of a particular stimuli through a set of common characteristics of previous stimulus.

- SCHULZ in [1] has defined the generalization as follow: *«The generalization is the ability to generate and act according to predictions for new observations based on underlying commonalities, i.e. is the ability to make predictive inferences about unobserved outcomes."*

Generalization may occur in many different form: inductive reasoning, concept learning, and word learning. In our work, we are interested on concept learning.

---

1 The Definitive Glossary of Higher Mathematical Jargon

3. Concept learning

3.1 Concept

A concept is a subset of objects having a common features, defined over a large set of objects [Example: The concept horse is the subset of all objects that belongs to the category of horses].

3.2 Learning

Learning is a central part of the cognitive process, it can be defined as: "The activity or process of gaining knowledge or skill by studying, practicing, being taught, or experiencing something" [1].

A good understanding of human's learning behavior is the stepping stone for understanding how can machine be able to learn as humans do, in the following are the various categories of human learning methods:

- **Rote learning (memorization):** Memorizing things without knowing the concept/ logic behind them.
- **Passive learning (instructions):** Learning from a teacher/expert.
- **Analogy (experience):** Learning new things from our past experience.
- **Inductive learning (experience):** On the basis of past experience formulating a generalized concept.
- **Deductive learning:** Deriving new facts from past facts.

3.2.1 Learning theories:

Learning theories can be classified to normative or descriptive [2]:

- **Normative theories:** provide a standard for optimal learning behavior.
- **Descriptive theories:** tend to build a good understanding of the actual human learning behavior.

3.2.2 Learning methods:

Learning methods are classified as data intensive or knowledge intensive [3].

- **Data *intensive* methods:**

Learning symbolic concepts using data intensive similarity based methods, the learner is provided with a large number of related examples and is asked to identify their similarities then generalize the concept embedded. There is two approaches of data intensive methods:

a. **The learning of heuristics represented as a collection of production rules:** in this approach learning is carried out by modifying each of the rules based on decisions made by these rules and on positive and negative examples.

b. **Decision theoretic methods:** is the second class of data intensive learning methods that use statistical decision theory to discriminate probability patterns exhibited in learning examples. Example of decision theoretic methods include evolutionary programming genetic algorithms, classifier systems and artificial neural networks.

- **Knowledge intensive methods:**

These methods rely on domain specific knowledge to learn and generalize, in explanation based learning, the learner uses domain knowledge to analyze a single training example and the target concept to produce a generalization of the example and deductive justification for this generalization. Knowledge intensive methods work well when the concept to be generalized can be deduced from the domain knowledge.

3.3 Concept learning

Inferring a general definition of a concept from a specific training examples, by searching for the hypothesis that is consistent with these examples over a predefined a space of potential hypotheses.

3.3.1 Computational approaches to concept learning

There exist two classes of different concept learning approaches differ in their ability to generalize reasonably from limited positive examples, depending to the way they model concepts [4].

- **Discriminative approaches:**

Do not enclose an explicit model to concepts, but follow a procedure for discriminating category members from non-members (eg. K-Nearest neighbor classification).

This approaches cannot learn to discriminate positive and negative instances if they have observed juts positive examples.

- **Distributional approaches:**

The concept in these approaches is modelled as a probability distribution over some feature space, a new instances X are classified as members of concept C if their probability p (X|C) exceeds certain threshold θ.

This class of approaches includes “novelty detection” techniques based on Bayesian nets and auto encoder networks.

3.3.2 Concept learning, hypotheses representation and search strategy

The choice of hypotheses representation and the search strategy are fundamentals for any concept learning algorithm.

- **Hypothesis space:**

  It is the set of all candidate hypotheses that may be the true extension for the concept to be learned in most of cases. The hypothesis space is determined by the designer of the learning algorithm.

  Concept learning can be viewed as the task of searching through a large space of hypotheses implicitly defined by the hypothesis representation. The goal of this search is to find the hypothesis that best fits the training examples.

- **Search strategy:**

  A central characteristic for any good concept learning algorithm is its ability to search exhaustively over a large or may be infinite hypotheses space, without the need of an explicate enumeration of every hypothesis. This can be accomplished by relying on the naturally occurring structure over the hypotheses space “genral_to_specific” hypothesis ordering [7].

- **General to specific hypotheses ordering:**

  Most of concept learning algorithms take advantage of naturally occurring structure of hypotheses what helps to organize the search over the hypotheses space and enables them to explore the hypotheses space exhaustively. In the following we will show how can it be used to define a sense of order between hypotheses.

  ✓ **Notes:**

  – For a given example x in the space of instances X, and a hypothesis h of the hypotheses space H: x satisfies h if and only if h(x)=1.

  – The relationships presented in the following are both defined depending on the instances that satisfy the hypotheses independently of the target concept (concept to be learned).

1. More general than relation $>_g$:

   Consider these hypotheses:

   h_1 = {Rainy, Warm, Strong}, h_2 = {Rainy,?, Strong}.

Obviously the hypothesis h_2 imposes less constraints on the instances than h_1, therefore it classifies more instances as positive, thus we say that h_2 is more general than h_1.

2. More general than or equal to relation $\geq_g$:

We say that a hypothesis h_j is more general than or equal to hypothesis h_k if and only if the instances that satisfy h_j ,satisfy h_k as well ($h_j \geq_g h_k$).

- **Concept learning algorithms based on general to specific hypotheses ordering:**

Many concept learning algorithms use the general to specific hypotheses ordering in order to make the search through the hypotheses space more efficient and practically feasible, following are some of these algorithms.

1. FIND S algorithm

The use of hypotheses ordering in FIND S algorithm resides in that it starts with the most specific possible hypothesis in H, then generalizes it each time it fails to classify positive training example as positive instance.

- **Algorithm**

1. Initialize h to the most specific hypothesis in the hypothesis space H.
2. For each positive example x

   For each attribute constraint a in h

   If the constraint a is satisfied by x then do nothing

   Else replace a in h by the next more general constraint that is satisfied by x
3. Output the hypothesis h.

- **Limits of FINDS algorithm:**

– There is no way to determine if it has found the only hypothesis in H consistent with the given data or there are others.

– Inconsistent sets of training examples (contains noise or errors) may mislead FIND S algorithm.

– In case there are multiple hypotheses consistent with the training examples, FIND-S will find the most specific.

2. Candidate elimination algorithm

Candidate elimination algorithm starts by initializing the version space to the set of all hypotheses in H, where G boundary set contains the most general hypothesis in H, and S boundary set contains the most specific hypothesis in H, the strength of this algorithm is in outputting a general description of the set of all hypotheses consistent with the training and this is achieved by basing on the general to specific hypothesis ordering.

- **Algorithm**

  • Initialize G to the set of maximally general hypotheses in H

  • Initialize S to the set of maximally specific hypotheses in H

  • For each training example d, do

  – If d is a positive example

  Remove from G any hypothesis inconsistent with d,

  For each hypothesis s in S that is not consistent with d,-

  Remove s from S

  Add to S all minimal generalizations h of s such that

  » h is consistent with d, and some member of G is more general than h

  Remove from S any hypothesis that is more general than another hypothesis in S

  – If d is a negative example

  Remove from S any hypothesis inconsistent with d

  For each hypothesis g in G that is not consistent with d

  Remove g from G

  Add to G all minimal specializations h of g such that

  » h is consistent with d, and some member of S is more specific than h

  Remove from G any hypothesis that is less general than another hypothesis in G.

- **Limits Candidate-Elimination Algorithm:**

  It is not robust to noisy data or to situations in which the unknown target concept is not expressible in the provided hypothesis space.

3. List then eliminate Algorithm:

This algorithm begins with a full version space**.** After observing every training example, the hypothesis that does not agree with the training example will be is eliminated, as more examples are considered the version space will narrow to remain only one hypothesis that is consistent with all observed examples.

- **Algorithm**

1. Version space ← a list containing every hypothesis in H.
2. For each training example

   Remove each hypothesis which is inconsistent with the training example
3. Output the list of hypotheses in the version space.

- **Limits of List-Then-Eliminate Algorithm:**
  - If the available data is insufficient to narrow the version space, all hypothesis will have same importance.
  - It requires enumerating all possible hypothesis in H, which is unrealistic for much of hypothesis spaces.

3.3.3 Concept learning theories (literature review)

Understanding the psychological theory of concept learning is the first step for understanding human concept learning, and this last constitutes of some basic theories as presented in the following:

- **Rule based theory:**

Under the rule based theory for concept learning, concepts are represented as rules to classify objects based on their features, the learning process is carried out by providing the learner with rules and training examples then is asked what degree of beliefs it gives to each rule. For example a radiologist needs to classify a suspicious spot au X-ray either as a tumor or as natural issue attention, rule based theories suggests that the radiologist observe if the specific properties of the X-ray meet the same criteria as tumor, this theory was used in many works like[24][25].

- **Exemplar theory:**

The exemplary theory argues that peoples categorize a new stimuli by comparing it with instances already stored in a memory, these instances are called ”Exemplar”, In this theory, it is hypothesized that learners store examples verbatim. For example, the model suggests that individuals create the category “dog” by memorizing a set of all doges they ever experienced: Dalmatian, poodle, etc. and when a new stimulus is enough similar to some of the stored dog examples they will categorize it to the “dog” category. Some researches [26] have been done about exemplar theory and it have been used in others [27]

- **Prototype theory:**

Prototype theory is similar to the exemplar theory in that they both assert the importance of similarity in categorization, but instead of using exemplar, it uses prototype which is an abstract average of the category members, this theory proposes that a new stimuli belongs to a given category if it has a high degree of resemblance with the prototype of this category. Prototype theory had been applied in [28][29].

- **Explanation theory:**

In explanation based theory of learning, the acquisition of a new concept is done by experiencing examples of it and extract the relevant information from it, in other word the learner observe an examples and explain what it learns from this example by forming a general understanding from this example, a typical example of explanation based learner in machine learning is a program that learns to play chess by examples. Explanation based theory can be viewed as a method that perform: generalization, chunking, operationalization and analogy. gave a deep view about this theory[30][31].

- **Bayesian theory:**

The Bayesian theory of learning is based on the Bayesian inference. Once assuming H, the hypothesis space of candidate extensions for the concept to be learned C, the Bayesian learner is provided with a set of N examples X={x_1,$x_{2,\dots,}$$x_n$},that are random samples from the true extension of the target concept C and then is asked to infer what other instances may fall under the consequence region of C, which is given by computing the probability of generalizing the learned concept C to new entity y as follows:

$$p(y \in C|X) = \frac{\sum_{h \in Hx,y} p(h).p(X|h)}{\sum_{h \in Hx} p(h).p(X|h)}$$

Before observing any example a prior probability distribution is associated to the hypothesis space **H**, reflecting the learner's contextual knowledge and background that he brings to the task. The a priori probability of a hypothesis **h** ∈ **H** is defined as an Erlang distribution as follows:

$$p(h) = \frac{|h|}{\delta^2} \exp(-\frac{|h|}{\delta})$$

After observing some examples the state of knowledge of the Bayesian learner, that represents its beliefs about which hypothesis may be the true extension of the concept C is represented by a posterior probability, computed using Bayes' rule combining the likelihood p(X|h) and prior probability p(h) .as follows :

$$p(h|X) = p(X|h).p(h)$$

The likelihood **p(X|h)** measures the probability of observing the examples **X** if **h** was the true extension of the target concept .under the strong sampling (the examples are random samples from the concept's extension) the likelihood is given by the following equation:

$$p(X|h) = \begin{cases} \left[\frac{1}{|h|}\right]^n & if\ X \in h \\ 0 & otherwise \end{cases}$$

Where **n** is the number of observed examples, n=|X|.

Bayesian learning performed well in different learning domains including medicine (symptoms and diseases) [21], bioinformatics (traits and genes) [22], and speech recognition [23] and so on. And it has been expanded to solve some problems effectively like in [17], a detailed explanation of this theory is presented in [2] [3]

- **Component display theory:**

Component display learning theory specifies how to design instructions for any cognitive domain, it separates the content from instructional strategy .Merrill in [7] classifies learning into two dimensions, **conten**t: consists of facts, concepts, procedures and principles. And **performance:** made up of remembering, using and generalities. Forming a matrix using these two dimensions, helps the instructor to determine which elements are the goals of the learner.it is a set up to determine the level of performance needed for an area of content.it have been used in many application domains.

4. Conclusion

Concept learning is a very vast term that can be taken from different points of view according to the search perspectives, in this chapter we tried to guide the reader by giving a clear idea about the context that we will be working on, as well provided a simple explanation to what it should be understood in this level, in the coming chapters we will cover the tool that we will be using in our work which is semantic networks then we will present our work and the obtained results.

Chapter
02

Semantic Networks

1. Introduction

Data is the basis of any project associated with artificial intelligence, especially in the areas of machine learning. The way this data is represented varies according to the project needs. Several knowledge representation techniques have been used in order to ease the research in this field, starting from logical predicate to production rules, frames and semantic networks (figure1). Semantic networks is the one among the mentioned representations that best fits our interest.

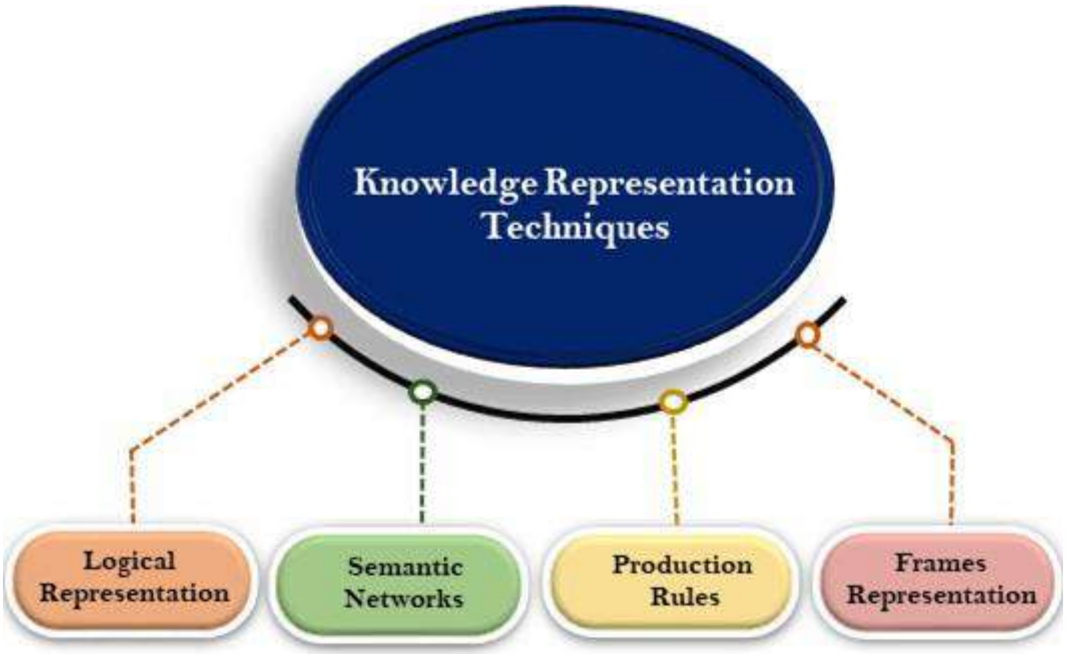


Figure 1: knowledge representation techniques

In the literature, several semantic networks have been proposed for different projects. Each of them has its own characteristics that distinguish it from the others (eg. Word Net, Concept Net…etc.). In our project we will work on the semantic network Concept Net.

This chapter includes two parts, in the first one, we review some of the basics of semantic sets whereas the second part gives a summary of Concept Net.

2. Semantic Networks

2.1 Definition

Several definitions have been attributed to the term semantic networks among that we can quote:

- Semantic networks is a type of data representation incorporating linguistic information that describes concepts or objects and the relation or dependency between them
- A semantic network is a graphic notation for representing knowledge in patterns of interconnected nodes Semantic networks require three constituent parts:

– A syntax that specifies the types of nodes knowledge that can be considered.
– Specification of the meaning or semantics those nodes or links and the entire network can represent.

Inference rules.

2.2 Semantic network representation

Semantic network can be presented in two different ways:

2.2.1 Graphical representation

In this way, the semantic network is represented as a graph structures that encode taxonomic knowledge of objects and their properties:

- Nodes labelled by relation constants corresponding to either taxonomic categories or properties
- Nodes labelled by object constants corresponding to objects in the domain
- There are three kinds of arcs connecting the nodes:

  Subset arcs (sometimes called *isa links*)

  Set membership arcs (sometimes called *instance links*)

- Example:

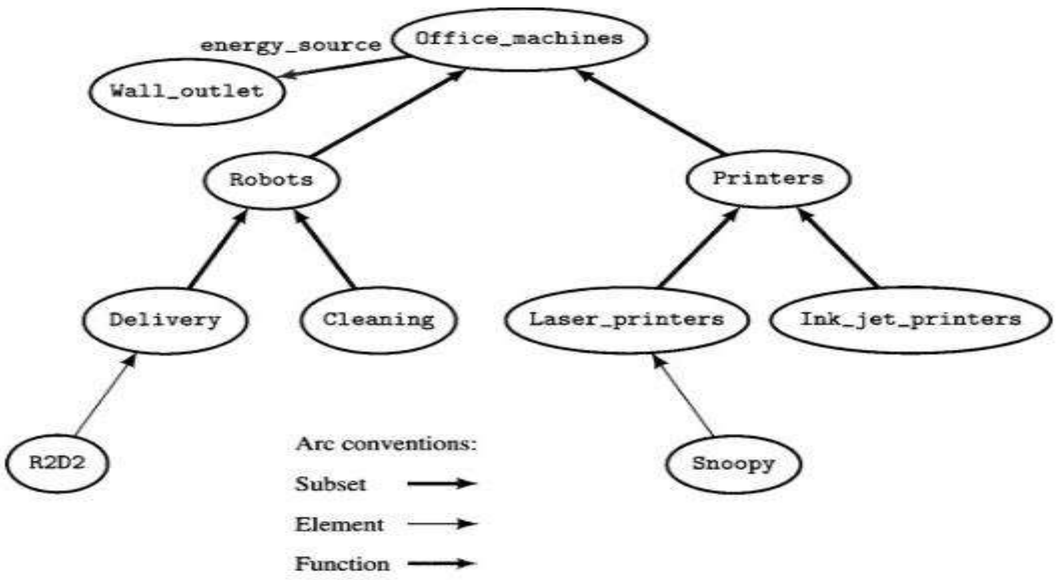
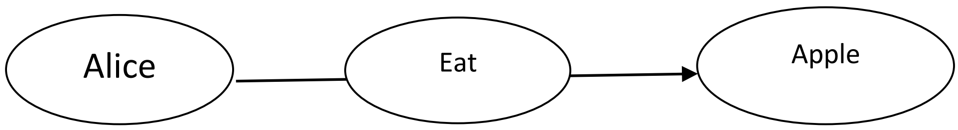


Figure 2: A simple example of semantic network representation

2.2.2 **Non graphical representation**

To illustrate this representation, we will take a portion of graphical representation example and convert it into this representation:



The non-graphical representation is: **(Alice, eat, apple)**

2.3 Main Components of semantic networks

The main components of any semantic networks are the following:

2.3.1 Lexical components

They are represented in the constituent elements of a network so that the physical objects we represent them in the form of labelled nodes, and the relationships between them are represented in the form of links containing labels that illustrate the type of relationship, for example (**figure3**) has 6 labelled nodes and 5 labelled links.

2.3.2 Structural components

It represents how the previous components (lexical components) are organized, as the nodes and links represent the form directed graphs and the labels are placed on each of them to clarify the meaning of those form directed graphs, for example (**fig3**) shows the general structure of the nodes and links mention in the last component. .

2.3.3 Semantic components

This type of component represents the semantic meaning of the network so that it differs according to the area represented by the network based on the labels of nodes and links. For example the semantic meaning of the (**fig3**) is about area of animals.

2.3.4 Procedural part

The procedural part relates a lot to the type of learning networks, so that in the latter, it is possible to **create, modify** or **delete** some links and nodes, for example in ANNs when we training the network we can **modify** the weights.

2.4 Example of semantic network

Statements:

- Jerry is a cat.
- Jerry is a mammal
- Jerry is owned by Riya.
- Jerry is brown colored.

- All Mammals are animal

A Semantic Network that can be constructed from these statements is in the following:

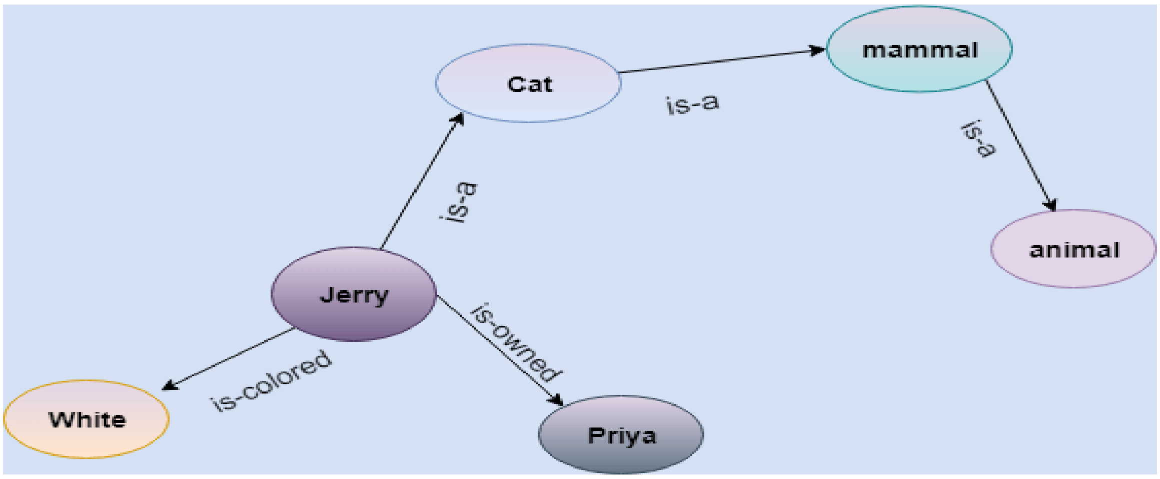


Figure 3: Simple Semantic Network

2.5 Semantic networks structure

The methods of representing the structure of semantic networks differ according to the way they are needed (the method of representing data). We will mention in this section some structural representations of semantic networks

2.5.1 A tree-structured hierarchy (Collins and Quillian (1969))

In this structure the concepts are represented as nodes and the additional nodes for characteristic attributes or predicates are linked to the most general level of the hierarchy to which they apply, and the connections between there determined by class-inclusion relations. This type of semantic network has some limitation. It is clearly appropriate only for certain taxonomically organized concepts, such as classes of animals or other natural kinds.

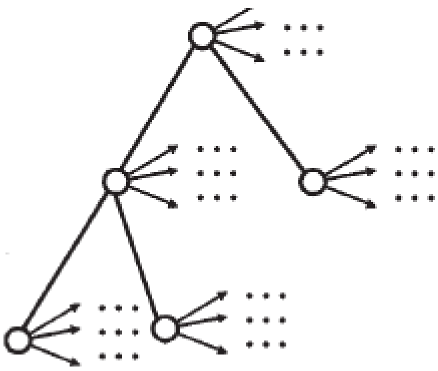


Figure 4: A tree-structured hierarchy (Collins and Quillian (1969))

2.5.2 An arbitrarily structured networks (Collins and Loftus)

In this structure each word or concept corresponding to a node and links between any two nodes that are directly associated in some way[12]. These types are essentially unstructured and are not characterized by any kind of large-scale structure.

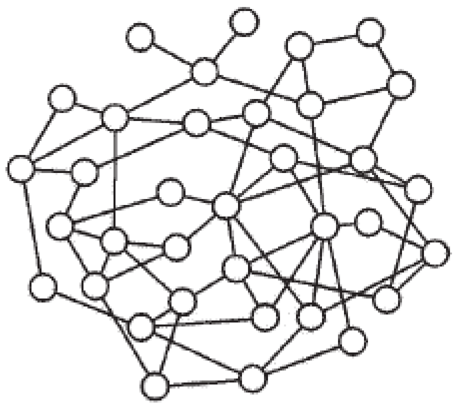


Figure 5: arbitrarily structured network

2.6 Use of semantic networks

The use of semantic network can emerge in different aspects:

- **Representing data:** it is used as a form of knowledge representation, for representing objects of the universe.
- **Revealing structure:** semantic networks provide an explicit representation of nodes, the relations between them, paths and all structure characteristics.
- **Supporting navigation:** semantic networks allow search and information retrieval and this is the main goal.

As well ,semantic networks are used in several domains of AI, including expert systems, natural language understanding , library document retrieval, analogical reasoning ,business planning, medical diagnosis, legal case analysis, robot control, deductive databases, intelligent Computer Aided Design, visual pattern recognition, simulated aircraft control, and many more.

2.7Types of semantic networks

The declarative graphic representation that is used to represent knowledge and support reasoning about it, is the common characteristic to all kinds of semantic networks, The following are six of the most common kinds of semantic networks [9]:

2.7.1 Definitional networks

Emphasize the *subtype* or *is-a* relation between a concept type and a newly defined subtype. The resulting network, also called a *generalization* or *sub sumption* hierarchy, supports the rule of *inheritance* for copying properties defined for a super type to all of its subtypes. Since definitions are true by definition, the information in these networks is often assumed to be necessarily true.

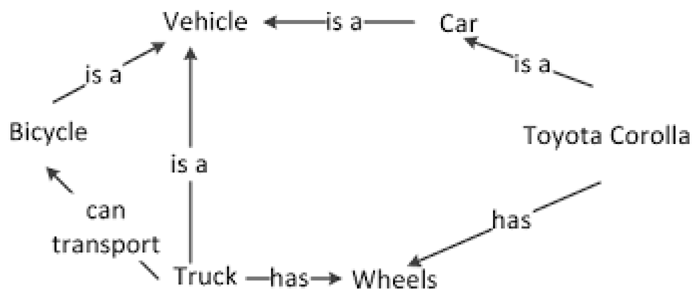


Figure 6: A definitional networks.

2.7.2 Assertional networks

Networks of these type are designed to assert propositions. Unlike definitional networks, the information in an assertional network is assumed to be contingently true, unless it is explicitly marked with a modal operator. Some assertional networks have been proposed as models of the *conceptual structures* underlying natural language semantics.

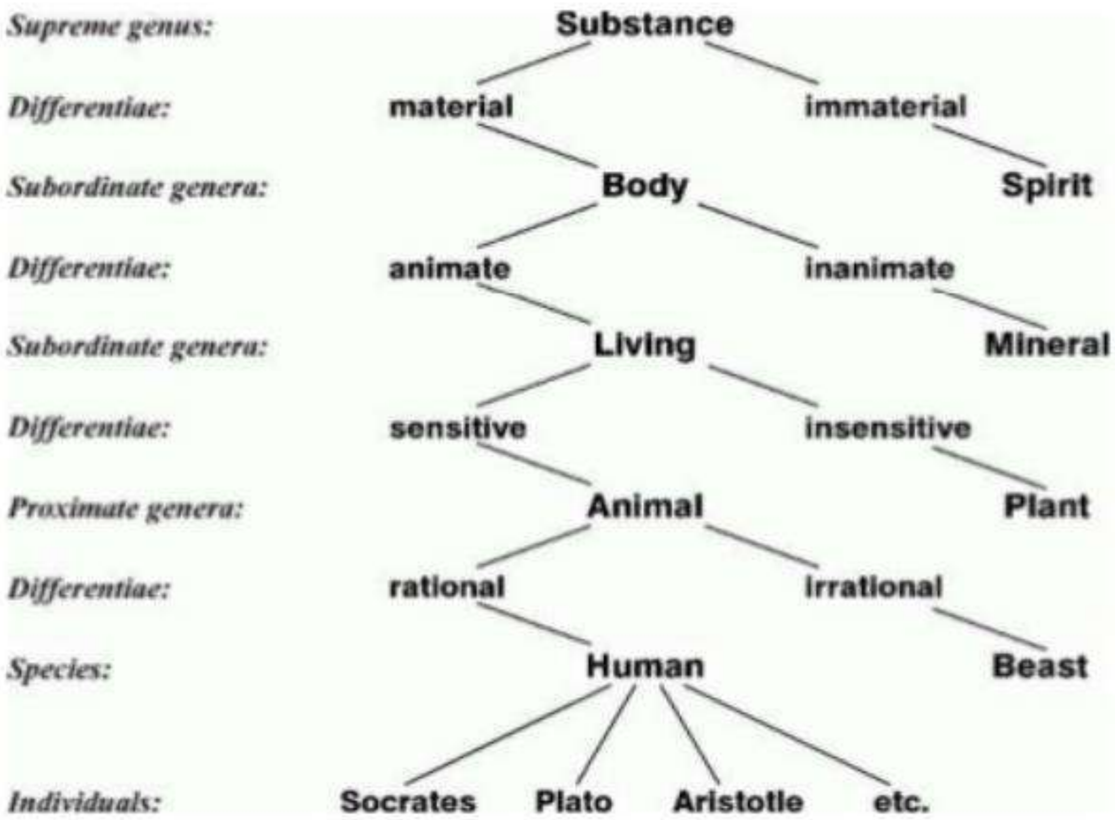


Figure 7: An assertional network

2.7.3 Implicational networks

Use implication as the primary relation for connecting nodes. They may be used to represent patterns of beliefs, causality, or inferences, For example we can represent the possible causes for slippery grass: like a box ,each box represent a proposition, and the arrows show the implications from one proposition to another.

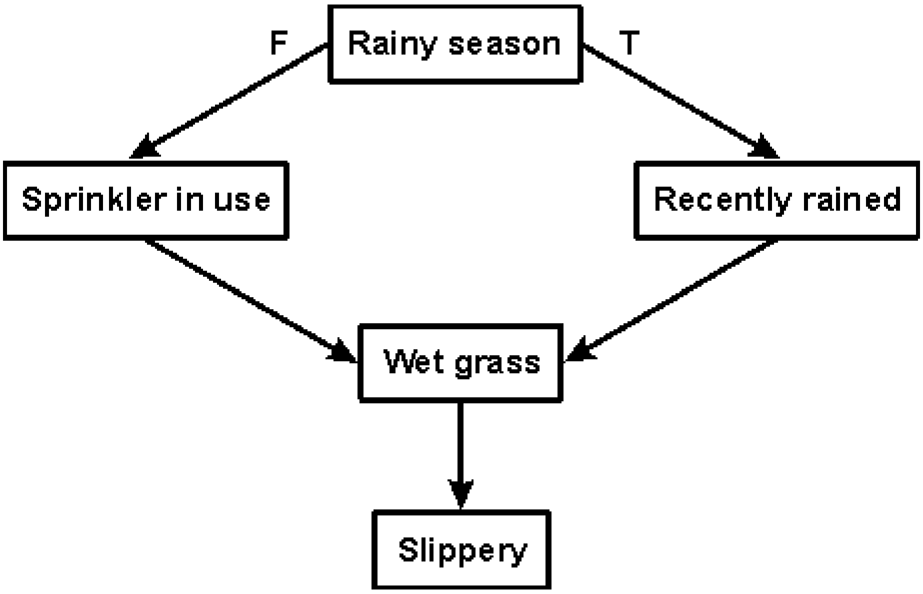


Figure 8: An implicational network for reasoning about wet grass

This figure represent this sentence:*" If it is the rainy season, the arrow marked T implies that it recently rained; if not, the arrow marked F implies that the sprinkler is in use. For boxes with only one outgoing arrow, the truth of the first proposition implies the truth of the second, but falsity of the first makes no prediction about the second"[9].*

2.7.4 Executable networks

Include some mechanism, such as marker passing or attached procedures, which can perform inferences, pass messages, or search for patterns and associations. For example we can represent this form X=(A+B)*S2N(C) like this:

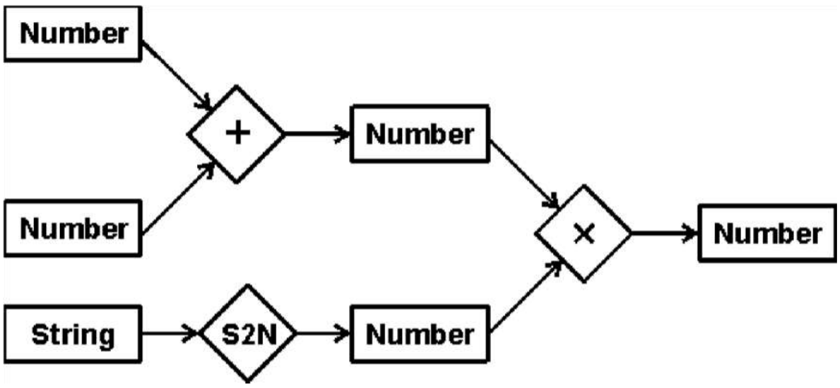


Figure 9: Learning networks

Build or extend their representations by acquiring knowledge from examples. The new knowledge may change the old network by adding and deleting nodes and arcs or by modifying numerical values, called *weights*, associated with the nodes and arcs,for example ANNs (Artifcial Neural Networks) see **fig 7**.

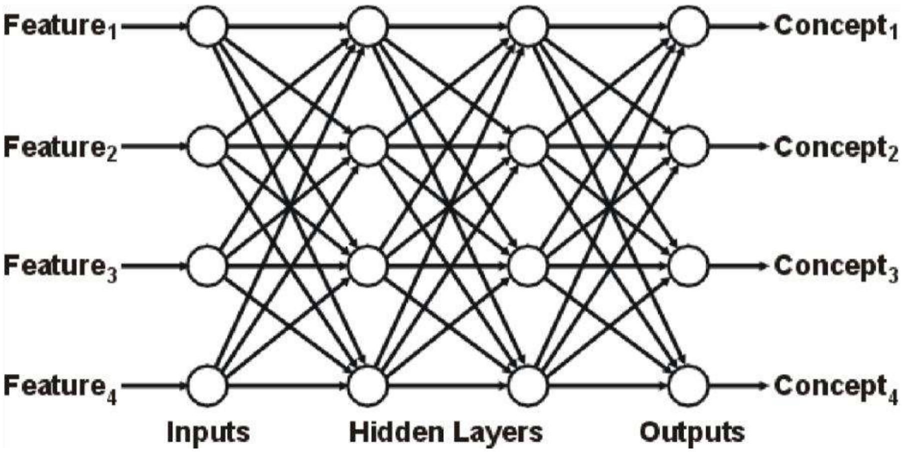


Figure 10: learning network (neural network)

2.7.5 Hybrid networks

Combine two or more of the previous techniques, either in a single network or in separate, but closely interacting networks.

2.8 Semantic network construction:

Semantic networks can be constructed either manually where they are annotated or maintained by domain experts(expensive process) or automatically by following a set of phases but it still a big challenge because there is no standardized way ,in the following is one of the most famous ways by which several strong semantic networks have been constructed like ConceptNet.

2.8.1 The phase of selecting the source of the information

In this part we will choose the source of data from which we can take the information so that the latter is equipped for that, meaning that the data is semi-strutted, and we mention one of these known sources, which is profitable. We mention an example of one of these known sources, which is Open Mind Common Sense(OMCS)corpus.

- OMCS:" *is a freely available crowd-sourced knowledge base of natural language statements about the world. The goal of OMCS is to provide intuition to AI systems and applications by giving them access to a broad collection of basic information and the computational tools to work with this data"[6]*

2.8.2 The phase of Extraction

In this phase we will defined the rule of extraction ,we do it in this steps:

- Set of extraction rules are used to map from English sentences of knowledge source into binary-relation assertions of semantic network (ex: cat is an animal -→ Isa (Cat, Animal)),and the sentences are no suitable relation-type may still be extracted into generic.
- Extraction rules are regular expression patterns crafted to exploit the already semi-structured nature of most sentences.
- Each sentence is given a surface parse by NLP (Natural Language Processing) so that syntactic and semantic constraints can be enforced on the nodes.
- As a result guaranteed syntactic structure, facilitating their computability.

2.8.3 phase of Normalization

"*Extracted nodes are also normalized. Errant spelling is corrected by an unsupervised spellchecker, and syntactic constructs (i.e. verbs, noun phrases, prepositional phrases, and adjectival phrases) are stripped of determiners (e.g. 'the' and 'a'), modals, and other semantically peripheral features. Words are stripped of tense (e.g. 'is/are/were '→'be') and number (e.g. 'apples '→'apple'), reducing them to a canonical 'lemma' form.* "[14]

2.8.4 The phase of relaxation

Meant to smooth over semantic gaps and to improve the connectivity of the network.

- Merged the duplicate assertion and add new metadata called 'frequency' to each predicate-relation to track how many times something is uttered (weights).
- Use the 'Is A' hierarchical relation to heuristically 'lift' knowledge from the children nodes to the parent node.
- Example:

  [(Is A 'apple' 'fruit');

  (Is A 'banana' 'fruit');

  (Is A 'peach' 'fruit')]

  AND

  [(Property Of 'apple' 'sweet');

  (Property Of 'banana' 'sweet');

  (Property Of 'peach' 'sweet')]

IMPLIES

(Property Of ‘fruit’ ‘sweet’).

- Relate more specific knowledge to more general knowledge this process called thematic and lexical generalizations.
- Example:

  (‘Buy food’, ‘buy’)

  (‘Purchase food’ ‘buy’)

- Lifted a noun phrase nodes contain adjectival modifiers and reified as additional Property Of knowledge.
- Example:

  [(Is A ‘apple’ ‘red round object’);

  (Is A ‘apple’ ‘red fruit’)]

  IMPLIES

  (Property Of ‘apple’ ‘red’).[5]

2.9 Advantages of using Semantic Nets

The most important advantages that distinguish semantic networks over their counterparts from other types of data representation is that they are simple and can be easily implemented and understood, And when we compare it with the logical representation, we find it more natural ,characterized by greater cognitive adequacy and They are has a greater expressiveness, in addition to that permit using of effective inference algorithm (graphical algorithm).

2.10 Limits of Semantic Nets

Since each representation has limits, we will mention some limits for semantic networks:

- There is no standard definition for link names
- Semantic Nets are not intelligent, dependent on the creator
- Undistinguished nodes that represent classes and that represents individual objects
- Links on object represent only binary relations
- Negation and disjunction and general taxonomical knowledge are not easily expressed.

3. Concept Net

3.1 Definition Of concept Net

Concept Net is a semantic network knowledge graph that connects word and phrases of natural language with labelled weights and edges (assertion), has a common sense knowledge that it is a part of AI, Concept Net is used for natural language understanding that describe general human knowledge and how it is in natural language[10].

"Is *a graph whose edges, or assertions, express common sense relationships between two short phrases, known as concepts . The edges are labeled from a defined set of relations, such as Is A, Has A, or Used For, expressing what relationship holds between the concepts. Each assertion additionally has a score to indicate its reliability, which increases either when a contributor votes for a statement through our Web site or when multiple contributors submit equivalent statements independently.*"[14].

3.2 ConceptNet Content

The information contained in ConceptNet includes relations between everyday objects ("Books are used for reading."), Information on people's priorities and goals ("People want to be respected.") and Affectual information ("Arguments make people angry.")[14]

3.3 Structure Of Concept Net

ConceptNet is conceptually represented as a hyper graph. Its assertions can be seen as edges that connect its nodes, these nodes are either concepts (words and phrases) or relations that are nodes as well [16].

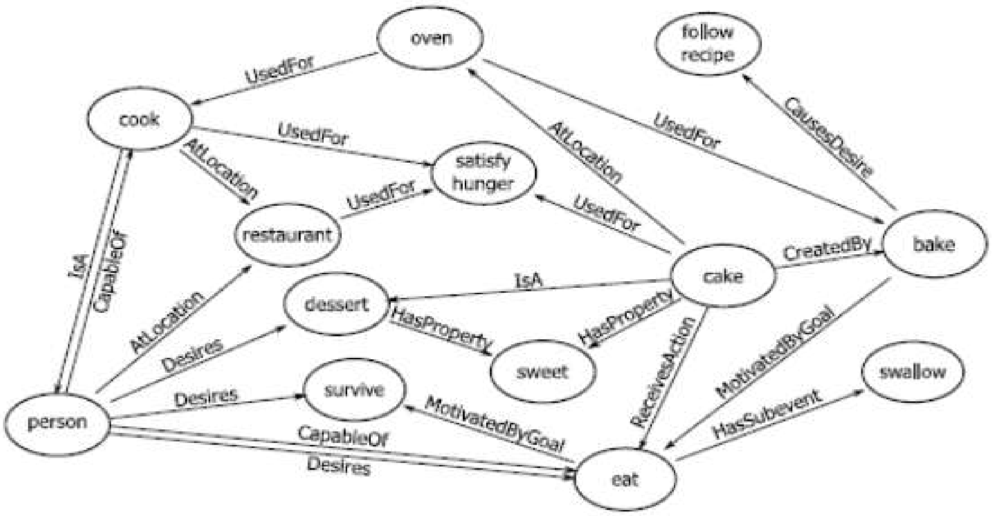


Figure 11: A high-level view of the knowledge Concept Net has about a cluster of related concepts

a. Node

Is a word or phrase of a natural language, often a common word in its disambiguated form. It is represented by the "URI /c/en/word or phrase "[15], for example, we can represent the node "person" like this "URI /c/en/person".

b. Relation

"ConceptNet uses a closed class of selected relations such as *Is A*, *Used For*, and *Capable Of*, intended to represent a relationship independently of the language or the source of the terms it connects."[15], it represented by" URL /r/name of relation" for example URL /r/Part Of"

- Symmetric Relation

Is all the relation the directionality of these edges is unimportant such as *Antonym ,Distinct From*, *Etymologically Related To*, *Located Near*, *Related To* , *Similar To* , and *Synonym*

- Asymmetric Relation

Is unlike symmetric relation such as *At Location*, *Capable Of*, *Causes*, *Causes Desire*, *Created By*, *Defined As*, *Derived From*, *Desires*, *Entails*, *External URL*, *Form Of*, *Has A*, *Has Context*, *Has First Sub event*, *Has Last Sub event*, *Has Prerequisite*, *Has Property*, *Instance Of*, Is A, *Made Of*, *Manner Of*, *Motivated By Goal*, *Obstructed By*, *Part Of*, *Receives Action*, *Sense Of*, *Symbol Of*, and *Used For.[15]*

In the table 1 below we maintain the most common Interlingua relations we identify in ConceptNet:

Relation	Sentence pattern	Relation	Sentence pattern
IsA	*NP* is a kind of *NP*.	LocatedNear	You are likely to find *NP* near *NP*.
UsedFor	*NP* is used for *VP*.	DefinedAs	*NP* is defined as *NP*.
HasA	*NP* has *NP*.	SymbolOf	*NP* represents *NP*.
CapableOf	*NP* can *VP*.	ReceivesAction	*NP* can be *VP*.
Desires	*NP* wants to *VP*.	HasPrerequisite	*NP*|*VP* requires *NP*|*VP*.
CreatedBy	You make *NP* by *VP*.	MotivatedByGoal	You would *VP* because you want *VP*.
PartOf	*NP* is part of *NP*.	CausesDesire	*NP* would make you want to *VP*.
Causes	The effect of *VP* is *NP*|*VP*.	MadeOf	*NP* is made of *NP*.
HasFirstSubevent	The first thing you do when you *VP* is *NP*|*VP*.	HasSubevent	One of the things you do when you *VP* is *NP*|*VP*.
AtLocation	Somewhere *NP* can be is *NP*.	HasLastSubevent	The last thing you do when you *VP* is *NP*|*VP*.
HasProperty	*NP* is *AP*.		

Table 1:The interlingua relations in Concept Net with example sentence frames in English

c. Linked Data

ConceptNet imports knowledge from some other systems, such as Word Net, into its own representation. These other systems have their own target vocabularies that need to be aligned with ConceptNet for imported these knowledge graph will be connected to ConceptNet nodes via the relation *External URL*, pointing to an absolute URL that represents that term in that external resource.

3.4Knowledge Sources

Concept Net contains a tremendous amount of information in all magazines and in all languages collected from many sources that we mention:

- The Open Mind Common Sense website
- Sister projects to OMCS in Portuguese
- The multilingual data, including translations between assertions, collected by Global Mind.
- Games with a purpose
- The English Wikitionary
- Word Net 3.0 including cross references to its RDF definition
- The semantic connections between Wikipedia articles represented in DBPedia [10]

3.5Versions Of concept Net

In this section we will mention all the current versions of ConceptNet and some characteristics of each of them:

- Concept Net 2: the first publicity version, it distributed as a python data structure, it was only in English.
- Concept Net 3: made into a SQL database, it include user interaction from web site, it were in English and Brazilian Portuguese.
- Concept Net 4: was quite similar to Concept Net 3, add a web API for accessing and querying data in concept Net 4.
- Concept Net 5: contains many representational improvements but the primary focus is to make the collection, storage, and querying of knowledge truly distributable

3.6 Motivation of Concept Net 5

ConceptNet version 5 contains a set of characteristics that made it much distinguished over other versions, starting with the addition of some languages such as Chinese, Portuguese, Japanese,,, ex. and the strength of their interconnection unlike in other version , as well as some sources, which is include the plan-oriented knowledge in Honda's Open Mind Indoor Common Sense , connections to knowledge in freebase , ontological connections to SUMO and MILO, in addition to that it supports the feature of extracting relational knowledge from web pages called machine reading such as Re Verb . And the most important thing that distinguishes it is that solve the problem of difficulty in accessing the data that was present in the other versions by separating the data from the interface.[16].

3.7 Comparison between ConceptNet and other representation (Word Net, Cycle)

ConceptNet provides a set of features not available in other knowledge representing such as Word Net or Cycle for example the concepts in concept Net are connected to natural language words and phrases that can be found in free text, not limited to a single language; they can be from any written language and focuses on a richer set of semantic relations unlike word Net.

The sources of knowledge in concept net are characterized by a high level of granularity and available through a common representation [10], In addition to that Concept Net's forte is contextual common sense reasoning unlike word net, Cyc and excels just at a lexical and a logical reasoning [13]

3.8 Some examples where we can use ConceptNet

Concept Net was used in many applications, ranging from building a system for analyzing the emotional content of text, to creating a dialog system for improving software specifications, recognizing activities of daily living, visualizing topics and trends in a corpus of unstructured text and to create public information displays by reading text about people and projects from a knowledge base [12].

Concept Net contains a lot of motivation what make it a source of interest for many scientific researchers to use in their projects and applications such as Common-sense ARIA, and recommendation system, Goose ,MAKEBELIEVE, Glo buddy[12].

3.9 Access of Concept Net

In order to use the huge amount of data that is contained within ConceptNet, that requires access to it either by downloading this data and using it as it was in previous versions of ConceptNet and this requires a large space because it has a huge statement and this is not suitable for those who want to use part of it only for example for a specific query. Therefore, the ConceptNet 5 plugins came to facilitate the process of accessing data through the Indexing feature.

We index ConceptNet 5 with a combination of Apache Solar and Mongo DB. We provide access to them through a REST API, for example when we need to know all the nodes related to concept ***Cat*** we can use just the API URl “http://api.conceptnet.io/related/c/en/**Cat** [10].

4. Conclusion

In this part, we presented a general overview about the semantic networks, then we discussed some kind of these networks such as ConceptNet, by highlighting the most important points that will enhance understanding the content of what we will present in our project as an important part of our work.

Chapter 03

Work Presentation

1. Introduction

In previous chapters we attempted to state the position of our contribution in machine learning literature, and point out what we will be using in our work .in this chapter we will first explain the principle of our algorithm then we will make an experimental evaluation by comparing between the performance of our algorithm with human generalization, then with the performance of the famous Bayesian frame work for generalization, finally we will discuss the results.

2. Work Principal:

The main goal of our work is to construct an algorithm that is capable of simulating human Concept learning. Human are able to learn concept from only limited number of examples because of their ability to generalize reasonably from a stimulus to others. Implementing this capability in machine seems feasible if we could find a base of knowledge that can take the place of human knowledge for machine and create a strategy to browse this base of knowledge and generalize effectively like human do. We chose Concept Net semantic network to be the base of knowledge of our algorithm, it is very convenient base of knowledge due to its richness. The search strategy that we proposed is based on two principals, inclusiveness and maximum semantic similarity value.in other word the algorithm will search for the concept that includes much of the concepts presented in the query, which has maximum semantic similarity value in relation to these concepts . To make the idea more clear, consider the query Q_1= {apple, apple, apple, cherry}, obviously the concept apple has more importance then the concept cherry because it is repeated many times, here we can conclude that the intended generalized concept is apple and thus the generalization given by our algorithm is apple (in the experimentation that we did some of the person were satisfied with this result and some wanted an upper level of generalization (eg. Fruit)).now consider the second query Q_2={apple ,cherry, strawberry}, all concepts of this query have same degree of importance and none of them is dominant ,thence the algorithm will search for the concept that includes most of these concepts ,from human perspective ,someone may say that they are all fruits, other may say red fruits and other one will say they all have red color ,all those possibilities are accepted by our algorithm adding some preferences based on semantic similarity value between concepts provided by concept net what gives priority to some concepts over others. In the case of Q_2,our algorithm will start searching in the first level in the taxonomy in relation to the concepts of the query, if there is no concept that includes all these concepts, we will select the N concepts that could include much of concepts ,save them and pass to the second level do same thing and pass to the next level if there is no general concept found .The algorithm stops if it finds a concept that includes all these concepts or it reaches certain depth without finding any general concept .if the algorithm could find a concept that includes all the concepts of the

query then this concept is the generalization otherwise there may be two reasons if there is no general concept found ,if the query contains a noise the algorithm will never find a concept that includes all the concepts of the query or because the structure of concept net cannot be exactly like human knowledge structure and it may miss some relations between concepts. The result given by our algorithm corresponding to the query Q_2 was: fruit with rank = 38.70, red with rank = 29.87, edible fruit with rank=13 and red fruit with rank =10.78.
in the fowling will give more detailed view of our algorithm and the evaluation we made.

3. Presentation of our algorithm:

3.1 Principle:

Our algorithm proceed in three main phases: primary phase, analyze phase and decision phase.

- Primary phase: in this phase we calculate the percentage (of contribution in formulating the query) to each concept in the query, the percentage of concept c from the query is given by

$$\text{Per}(c) = \frac{\text{rank of c}}{\text{length of the query}}$$

Where the rank of c is its number of occurrence in the query.

- Analyze phase: in this phase we will check if the generalization is included in the query, if there is a concept c with a percentage ≥0.7 then c will be the generalized concept, or if there is more than one concept such that their percentages are ≥0.4 the generalization will be these concepts, otherwise the algorithm will search in the provided knowledge space for the concept that includes all or most of the query concepts in its neighbors.
- Decision phase: once the algorithm stops the search, the elected concepts will be arranged in descending order by percentage then rank, the k first concepts include the generalization of the query.

3.2 Algorithm:

Begin:

1. List of concepts ←query
2. For each concept C in the query :

   Rank of C=number of occurrence of C in the query.

   Percentage of C=rank of C /length of the query.
3. If generalization is in the query then:
   - Return the generalized concept(s)
4. Else :
   - For each concept C in the list of concepts

For each neighbor NC of C:

Add NC to list of neighbors of query concepts.

Rank of NC=rank of C+ sum of weights of all the relations between C and NC.

- For each concept CN in the list of neighbors of query concepts:

  N_occ=0

  For each concept C of the query:

  If list of neighbors of C contains CN then:

  N_occ=N_occ+1

  Percentage of CN= $\frac{N\_occ}{\text{Number of concepts of the query}}$

- If there is a concept(s) with percentage=1.0 then:

  Arrange concepts descending by percentage and rank.

  Return k first concepts.

- Else save concepts with maximum percentage.
- If number of iterations < NI then:

  List of concepts ←list of neighbors of query concepts.

  Repeat from 4.

- Else:

  Arrange saved concepts descending by percentage and rank.

  Return k first concepts.

End

5. Experimental evaluation

The evaluation of our algorithm consists of two scenarios, in the first one we compared its performance with human generalization, and in the second scenario we compared it with the performance of Bayesian framework for generalization.

We used **top N** accuracy measure, which is the accuracy where the true class matches with one of the N most probable classes predicted by the model. And it is an appropriate measure for our work.

5.1 First scenario:

The purpose of this experimentation is to show how much could our algorithm perform closer to human. We formulated about 80 query from different domains using concept net data and compared the results given by our algorithm with human generalization in two ways:

- **First experimentation:**

We precise human intention for each query and compare it with the prediction of our algorithm, some of the obtained results are illustrated in the following:

Query	Human intention	Given prediction		Query	Human intention	Given prediction
2 potato, 2 tomato, 1 pepper	Vegetables	**vegetable** rank ---> 20.37 percentage ---> 1.0 **solanaceous_vegetable** rank ---> 11.0 percentage ---> 1.0		2 cygnets, 1 cob, 1whooper	Swan	**Swan** rank ---> 20.0 percentage ---> 1.0
2 apple, 2 cherry, 1 strawberry	Red fruits	**fruit** rank ---> 38.70 percentage ---> 1.0 **red** rank ---> 29.87 percentage ---> 1.0 **edible fruit** rank ---> 13.0 percentage ---> 1.0 **red fruit** rank ---> 10.78 percentage ---> 1.0		1cygnet, 1 seabird, gallinule	Aquatic bird	**bird** rank ---> 10.0 percentage ->0.65 **aquatic bird** rank ---> 6.0 percentage ->0.66
Plane ticket ,plane ,baggage	travel	**travel** rank ---> 13.82 percentage ---> 1.0 **airport** rank ---> 6.46 percentage ---0.66		Salat, hajj, sawm ,zakat ,shahada	Pillar of Islam	**Islam** rank ---> 14.0 percentage ---> 1.0 **pillar of Islam** rank ---> 12.0 percentage ---> 0.8

Table 2: some of the results given by our algorithm

We calculated the accuracy of our algorithm compared to human intention the results are in the following:

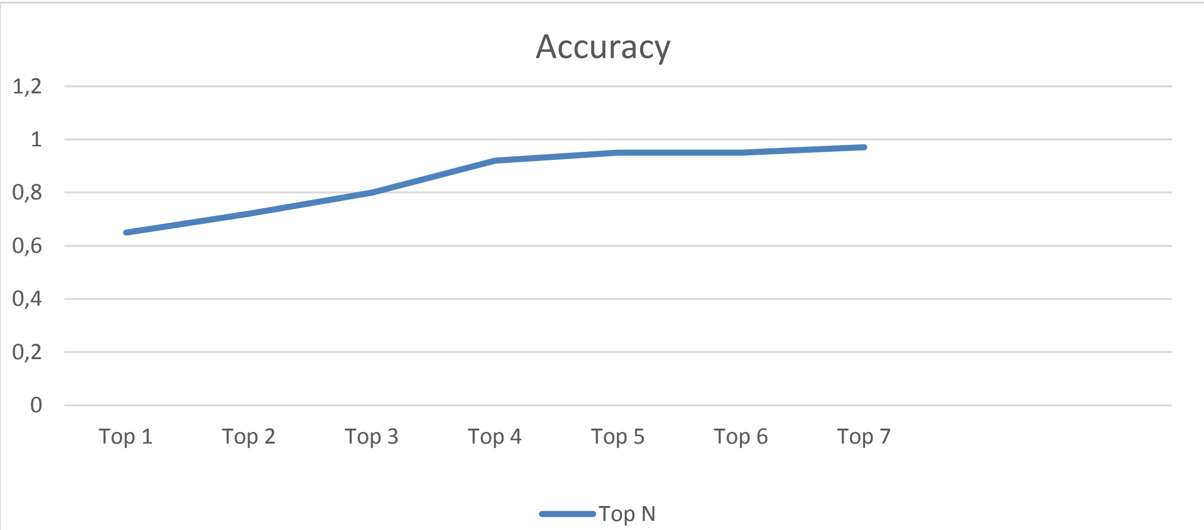


Figure 12: Accuracy of our algorithm compared to human intention

- **Second experimentation:**

In order to strengthen the evaluation of our algorithm compared to human generalization, we did an online questionnaire as shown in **figure 13**, where we invited more than 10 persons from different ages and categories, we provided them with some of the queries that were used previously and asked them about their degree of satisfaction (very good, good, weak or bad) with the result (first prediction) given by the algorithm for each query. The results are presented in **figure14**.

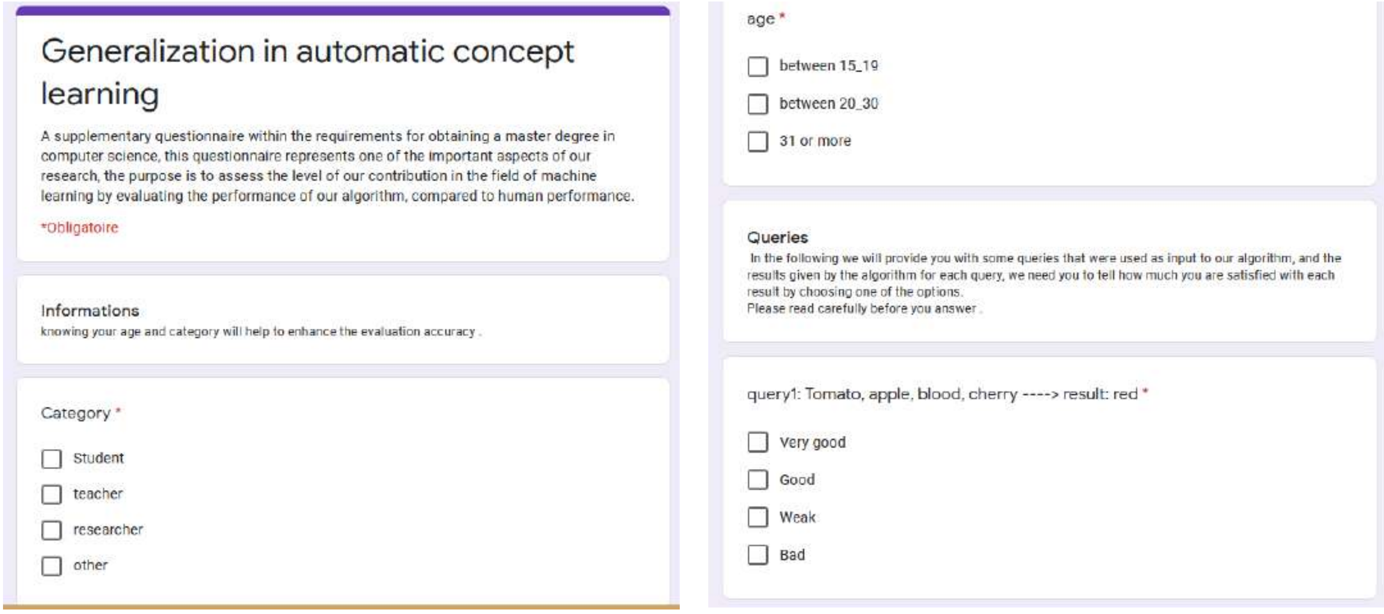


Figure 13: screenshot for part of the online questionnaire

Figure 14: human satisfaction with the prediction of our algorithm

As shown in the table of results the predictions of our algorithm matched human intention for all the queries, notice that the correct prediction was not always in first ranking because as we said before that the algorithm is based on two principals: **1.calculating** the value of **semantic similarity** between the generalized concept and the concepts of the query, this will give a high ranking to some of the concepts that have same degree of inclusiveness of query concepts over the others. **2.**the degree of **inclusiveness** ,in some cases because the intended concept does not include all the concepts presented in the query (in concept Net)thus other concepts that include more concepts of the query will have more priority, this is summarized in the accuracy of our algorithm (**figure 14**) the much we expand the scope of accepted predictions the much the accuracy increases .this cannot be counted as a limit for the algorithm because in second experimentation we noticed that even peoples had different degree of satisfaction with same query ,thus we can say that the performance of our algorithm was so satisfying and this is justified by the richness of concept net ,as well the exhaustive search strategy of our algorithm.

5.2 Second scenario:

We formulated about 20 query to compare the performance of our algorithm with the performance of the Bayesian framework for generalization, we calculated the accuracy to each approach the results are in the following:

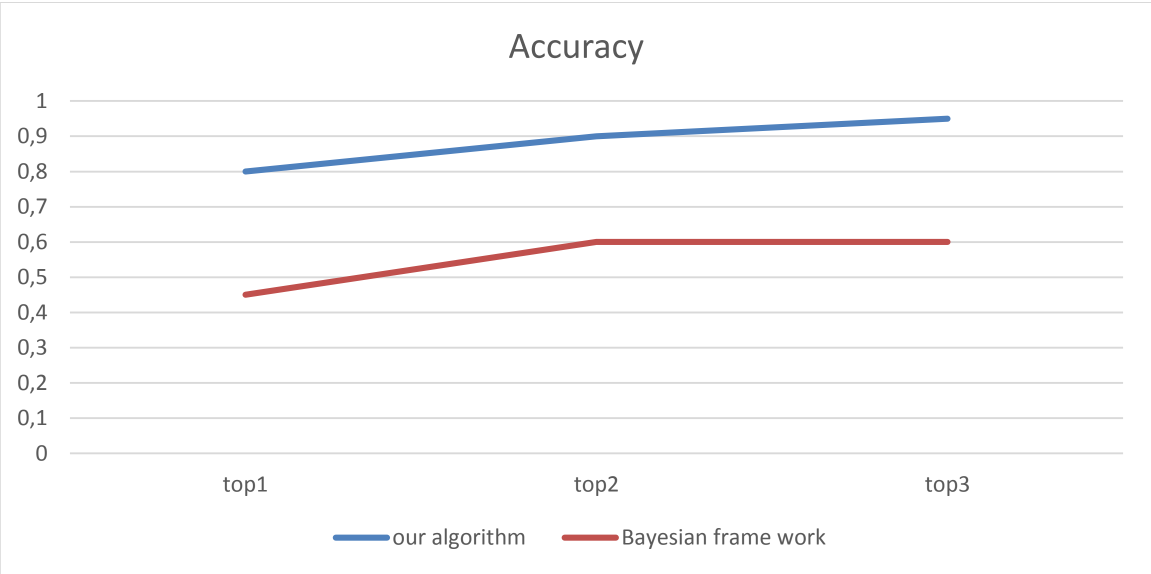


Figure 15: Accuracy of our algorithm and Bayesian framework

From the results presented above we can obviously see that our algorithm outperformed Bayesian framework with a significant difference. This is due to the flexibility of our algorithm with the content of queries and the tolerance of noise where in contrast Bayesian framework always generalize to the first concept it found to include all concepts of the query otherwise it will not return any result and this is not practical in most of the important problems encountered in the field of machine learning.

6. Conclusion:

In this chapter we explained the principal of our algorithm, presented the different experimental scenarios that we used to evaluate the performance of our algorithm and discussed the obtained results. The results were so satisfying compared to human generalization and surpassed Bayesian frame work performance.

General conclusion:

In machine learning literature several approaches have been proposed for automatic concept learning ,however one common problem of these approaches is that they all require a big mass of data in order to train machine effectively, but collecting sufficient data for every single domain of real world is such tiring and tedious task and unrealistic sometimes , This obstacle push researchers in artificial intelligence field to try to understand the mystery underling human capacity to learn concepts effectively from only few number of examples ,cognitive scientists gave a clear answer to this question ,they cast the problem of concept learning from few examples as a generalization problem, in this context many strategies have been suggested ,The Bayesian framework for automatic concept learning is the one that achieved a considerable success in many learning tasks, however it is only successful when the stimuli are represented in an appropriate psychological space. Our objective was to provide a new concept learning algorithm that would solve the problems encountered in previous works .the work was based on the use of semantic network, this later offered very rich learning space what helped the search strategy of our algorithm, the obtained results were so satisfying, our algorithm could achieve an accuracy of 95% compared to human generalization and outperformed Bayesian generalization with accuracy difference of 30%. This contribution open the door for coming contributions in this subject ,in future work we may use counter examples in order to narrow the possibilities and add more preference when choosing the generalized concept.

References

[1] Bharti, G. (2017, August 16). Retrieved from Knoldus: https://blog.knoldus.com/concept-learning-stepping-stone-towards-machine-learning/

[2] TENENBAUM, Joshua Brett. *A Bayesian framework for concept learning*. 1999. Thèse de doctorat. Massachusetts Institute of Technology.

[3] TENENBAUM, Joshua B. Bayesian modeling of human concept learning. In : *Advances in neural information processing systems*. 1999. p. 59-68.

[5] SHEPARD, Roger N. Toward a universal law of generalization for psychological science. *Science*, 1987, vol. 237, no 4820, p. 1317-1323.

[6] AUSTERWEIL, Joseph L., SANBORN, Sophia, et GRIFFITHS, Thomas L. Learning How to Generalize. *Cognitive science*, 2019, vol. 43, no 8, p. e12777.

[7] MITCHELL, Tom M. Does machine learning really work?. *AI magazine*, 1997, vol. 18, no 3, p. 11-11.

[8] MERRILL, M. David. Component display theory. *Instructional-design theories and models: An overview of their current status*, 1983, vol. 1, p. 282-333.

[9]: SOWA, John F. Semantic networks. *John_Florian_Sowa isi [2012-04-20 16: 51]> Author [2012-04-20 16: 51]*, 2012.

[10]: SPEER, Robert et HAVASI, Catherine. Representing General Relational Knowledge in ConceptNet 5. In : *LREC*. 2012. p. 3679-3686.

[11]:https://www.sciencedirect.com/topics/computer-science/semantic-network13:05 19/04/2020.

[12]: STEYVERS, Mark et TENENBAUM, Joshua B. The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive science*, 2005, vol. 29, no 1, p. 41-78.

[13]: LIU, Hugo et SINGH, Push. ConceptNet—a practical commonsense reasoning tool-kit. *BT technology journal*, 2004, vol. 22, no 4, p. 211-226.

[14]: HAVASI, Catherine, SPEER, Robert, ARNOLD, Kenneth, *et al.* Open mind common sense: Crowd-sourcing for common sense. In : *Proceedings of the 2Nd AAAI Conference on Collaboratively-Built Knowledge Sources and Artificial Intelligence*. 2010. p. 53-53.

[15]: SPEER, Robyn, CHIN, Joshua, et HAVASI, Catherine. Conceptnet 5.5: An open multilingual graph of general knowledge. *arXiv preprint arXiv:1612.03975*, 2016.

[16]: SPEER, Robert et HAVASI, Catherine. ConceptNet 5: A large semantic network for relational knowledge. In : *The People's Web Meets NLP*. Springer, Berlin, Heidelberg, 2013. p. 161-176.

[17] KORICHI, Meriem, KHERFI, Mohamed Lamine, BATOUCHE, Mohamed, *et al.* Extended Bayesian generalization model for understanding user's intention in semantics based images retrieval. *Multimedia Tools and Applications*, 2018, vol. 77, no 23, p. 31115-31138.

[18] JOHNSTONE, Michael N. et WOODWARD, Andrew. Towards effective algorithms for intelligent defense systems. In : *International Symposium on Cyberspace Safety and Security*. Springer, Berlin, Heidelberg, 2012. p. 498-508.

[19] AMIN, Javaria, SHARIF, Muhammad, RAZA, Mudassar, *et al.* Brain tumor detection using statistical and machine learning method. *Computer methods and programs in biomedicine*, 2019, vol. 177, p. 69-79.

[20] KIM, Dongil, KANG, Pilsung, CHO, Sungzoon, *et al.* Machine learning-based novelty detection for faulty wafer detection in semiconductor manufacturing. *Expert Systems with Applications*, 2012, vol. 39, no 4, p. 4075-4083.

[21] HENRIQUEZ, Ronald R. et KORPI-STEINER, Nichole. Bayesian inference dilemma in medical decision-making: A need for user-friendly probabilistic reasoning tools. *Clinical Chemistry*, 2016, vol. 62, no 9, p. 1285-1286.

[22] LIU, Jun S.. et LAWRENCE, Charles E.. . Bayesian inference on biopolymer models. *Bioinformatics (Oxford, England)*, 1999, vol. 15, no 1, p. 38-52.

[23] NORRIS, Dennis, MCQUEEN, James M., et CUTLER, Anne. Prediction, Bayesian inference and feedback in speech recognition. *Language, cognition and neuroscience*, 2016, vol. 31, no 1, p. 4-18

[24] Goodman, N. D., Tenenbaum, J. B., Feldman, J., & Griffiths, T. L. (2008). A rational analysis of rule-based concept learning. *Cognitive science*, *32*(1), 108-154.

[25] Hekanaho, J. (1997, May). GA-Based Rule Enhancement in Concept Learning. In *KDD* (pp. 183-186).

[26] Murphy, G. L. (2016). Is there an exemplar theory of concepts?. *Psychonomic Bulletin & Review*, *23*(4), 1035-1042.

[27] Ashby, F. G., & Rosedahl, L. (2017). A neural interpretation of exemplar theory. *Psychological review*, *124*(4), 472.

[28] Feldman, J. (2000). Minimization of Boolean complexity in human concept learning. *Nature*, *407*(6804), 630-633.

[29] Lawry, J., & Tang, Y. (2009). Uncertainty modelling for vague concepts: A prototype theory approach. *Artificial Intelligence*, *173*(18), 1539-1558.

[30] Flann, N. S., & Dietterich, T. G. (1989). A study of explanation-based methods for inductive learning. *Machine learning*, *4*(2), 187-226.
[31] Etzioni, O. (1993). A structural theory of explanation-based learning. *Artificial Intelligence*, *60*(1), 93-139.