Using Similarity Measures to Evaluate Biomedical Ontology Records

Osama Bassam J. Rabie
Department of Information Systems, Faculty of computers and information technology
King Abdulaziz University
Jeddah, Kingdom of Saudi Arabia
obrabie@kau.edu.sa

***Abstract*— In this paper, five binary similarity and distance measures are picked based on their usage in the biomedical field. The measures are applied on seven ontology records to evaluate the usage accuracy of relationships *part_of* and *located_in*. The accuracy is checked against the definition found on the online medical dictionary provided by the U.S. National Library of Medicine. Each measure is applied on each ontology records. The main contribution of this measure is to provide a mathematical prospective on the issues facing *part_of* relationship.**

Keywords—Biomedical Ontology; Similarity Measure; Operations Management

I. Introducing Motivation

The main motive behind this work is to contribute in resolving the issues facing ontologies. The interest in maintaining ontologies came from the fact that the amount of data exchanged on the internet is huge and without a mechanism to process (e.g. using ontologies) humans will not be able to process that amount of data. Processing the data on the internet can supply humans with useful information. The vision is to enable the web to process itself and provide us with useful information by using ontologies [4].

Measuring the inaccuracy of relationships used in biomedical ontologies contributes into achieving the semantic web vision. Several scholars indicate it is impossible to manage something without measuring it. To manage time, one should be able to count (or measure) the amount of time spent for each aspect of daily, monthly, and yearly activities to be able to decide if the time is spent in consideration of his or her priorities. Having the ability to measure is essential for the evaluation of any management approach. In this paper, an evaluation is conducted to measure the similarity (or distance) level between two relationships in the biomedical ontologies, the relationships *part_of* and *located_in*, under the consideration of definitions from MedlinePlus. MedlinePlus is an online medical dictionary offered by the U.S. National Library of Medicine, National Institutes of Health and Merriam-Webster, Incorporated.

The importance of biomedical ontologies is encouraging to write this paper. Ontologies in general are used to help in solving the challenge of having a huge amount of data to be processed. Ontologies creation is to enhance semantic information processing and knowledge exchange between machines. The biomedical ontologies are ontologies used in the field of bioinformatics [20]. As ontologies, the biomedical ontologies also deal with heterogeneous clinical data [26]. Two factors suggest the usage of biomedical ontologies. First, the biomedical ontologies are considered the most successful and practical innovation in the field of biomedicine [20]. Second, the biomedical ontologies considered the most essential and advanced application of the semantic web [10].

The work of this paper is mainly with the relationships *part_of* and *located_in*. The relationships *part_of* and its inverse *has_part* are the second most important structural relationships after the relationship *is_a* [24]. However, the relationship *part_of* can be confused with the relationship *located_in* [24]. In this paper, five similarity and distance measures are used to evaluate the usage accuracy of the relationship part-of at seven biomedical ontology records. The study contributes in proving the inconsistency of used biomedical ontology by evaluating the similarity of the used relationship and the dictionary definition.

A. *Ontology*

“An ontology is a "domain-specific dictionary", It captures the semantic meaning and relationship of terms which allows for further usage of the term’s concept [19]” [22]. In addition, it contains terms and their defined properties that can be executed by a computer. Ontologies are created by anyone with suitable tools. Ontology is a formal conceptual representation of domain concepts. "Ontology is a representation of universals; it describes what is general reality, not what is particular” [19]. Databases represent entities instances where ontologies represent entities classes. “Ontology provides engineers with the semantics of data which can be used with problem-solving methods along with reasoning services to produce a great system with fewer resources (Schulz et al. 2006)” [22]. Based on the graph structure of the ontology, having one inaccurate instance means the rest of the information obtained is going to be inaccurate. “Semantics enable machines to understand context and the type of data they

Sponsor is King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia.

utilize" [22]. Semantics give machines the ability to become closer to human understanding level [14]. Machines should be able to manage knowledge independently by using semantics [6].

"Currently, vast amounts of data require transmission and manipulation, and without sharing it, the data may become useless" [22]. Huge amount of data prevents people from being able to manually do transmission, interpretation, and manipulation. Semantic web (web 3.0) is meant to address those challenges [4, 18]. On the other hand, handling a large amount of heterogeneous data can be an integration challenging. "Ontology can solve this problem by providing computers with an understanding of information, and using it for reasoning without interference from humans" [22]. However, using ontologies can be challenging given that ontologies are written in languages still under development [18]. In addition, many ontologies including ontologies in use are corrupted, and finding an ontology to produce sufficient inference may be a challenge. Ontologies description is loos compared with databases [2] despite the effectiveness of using ontologies to process and exchange knowledge compared to the databases.

Ontology are important part of artificial intelligence implementation [22]. Software agents understand the conceptual representation of a specific domain knowledge embedded in ontologies. "Ontology provides information management systems with a way to handle unstructured contents, which may be impossible for computers to handle without ontologies" [22]. They provide knowledge description, natural language processing, and a reference for standardizing language modeling. Given the nature of ontology development, many ontologies co-exist with other ontologies for the same domain. Many ontologies are created without consulting a domain expert, which is a reason for having imprecise knowledge represented by those ontologies.

B. Paper Map

The binary similarity and distance measures used in the experiment are discussed in section 2. A brief description of the ontology used in the experiment is provided in section 3. In section 4, the calculation and discussion about them are presented. The recommendations and the ranking of the measures used shown in section 5. Finally, an appendix includes more details about some concepts used in the paper.

II. Similarity and Distance Measures Used

Binary Similarity Measures and Binary Distance Measures are measures used to determine the level of similarity between two (i.e. binary) patterns, or the level of dissimilarity or distance between the patterns [5, 8]. To apply the measures data should be converted into pairs of time series. In each pair, one time series represents the item to be compared where the other time series will represent the other item [5]. In addition, both time series should have the same length [5]. There are many applications for these measures, enhancing clustering [3], enhancing image-processing [28], and enhancing the detection of tumor [11]. Moreover, several fields, like ecology [16], biometrics [29], and handwriting recognition [7], are using the measures to help them deciding on the level of similarity between two patterns.

In this paper, the measures evaluate the accuracy of part_of usage in seven randomly selected biomedical ontology records. The rank of the binary similarity and distance measures is based on how close their similarity score to the average score of the ontology against medical dictionary (see table 1). Although the data can be represented in a variety of ways [9], in this study the binary data is the type used to be able to process more records in future publications.

The measures picked based on their potential importance in previous biomedical related studies. Being used in previous studies makes this study the first one to study the biomedical ontology by using measures used in biomedical studies. The fact that those measures were used in the field makes it more justifiable to use in studying the biomedical ontology. The following are the measures used and a brief justification for each measure of its candidacy to be considered in the field of biomedical and this study:

A. Jaccard [15]

This similarity measure is considered the first to be used in a biomedical related field, ecology [15], and started the idea of using similarity measures in different fields [8]. Jaccard is used in several ecological species classification [8, 15, 16] and in several biomedical ontology evaluations [23]. In addition, no large data is required for it to work properly [1]. Therefore, Jaccard is included in our experiment.

B. Ochiai [21]

Also known as Ochiai-I [8] where some literature consider it the same as the Cosine similarity [23, 29]. However, we consider Ochiai-I and Cosine similarity to be two different measures. Ochiai is a similarity measure used in many aspects related to the biomedical field especially in the field of biology [16, 17, 21, 23, 27, 29]. Therefore, we included Ochiai-I in our experiment.

C. Forbes [13]

Also known as ForbesI [8] was proposed to help in spices classification [12, 13]. Therefore, ForbesI is on our list.

D. Simpson [25]

It was proposed to study classification of faunal [25]. In addition, the measurement was used to estimate the semantic similarity in the biomedical domain [23]. It was used in studying the similarity of molecular fingerprints [28]. Therefore, Simpson is on the similarity measures list.

E. Yule [30]

Also known as YuleQ [8]. Yule was used in species classification [16]. In addition, It was used in studying the similarity of molecular fingerprints [28]. Therefore, it is on the list.

III. Biomedical Ontologies Used

The ontology used is called Foundation Model of Anatomy which is developed by Unified Medical Language System (UMLS). The ontology created by Unified Medical Language System (UMLS). The records picked from the Foundational Model of Anatomy(FMA) view of neuroanatomey ontology. The ontology version is 3.0 and the view upload date is Sep 18, 2009. The ontology category is anatomy.

There are three reasons for choosing this ontology. First, the projects using this ontology include:

- Electrophysiology Ontology (for The Johns Hopkins University).
- Cell line ontology (for Molecular Connections Pvt. Ltd).
- Influenza Ontology (for UT Southwestern, Univ of Maryland, and MITRE).
- NCBO Resource Index (for Stanford University).
- Plant Ontology (for Oregon State University).
- OntoCAT (for EMBL-EBI and University of Groningen).
- LogMap (for University of Oxford).
- SKELETOME (for School of ITEE, The University of Queensland).
- Neural ElectroMagnetic Ontologies (NEMO) (for University of Oregon and Georgia State University).
- Nomenclature for Anatomic Pathology (NAP) (for SIAPEC-IAP).
- OntoMaton (for Oxford University).
- BioAssay Ontology (for University of Miami).
- SemGen (for University of Washington).
- NCBO Annotator (for Stanford University).
- Adverse Event Reporting Ontology (AERO).

Second, Unified Medical Language System (UMLS) develop it. UMLS was made by U.S. National Library of Medicine (NLM) and still maintained by the NLM. According to NLM, NLM is “the world's largest medical library.” The UMLS project started beck in the 1986 by Donald A. B. Lindberg, M.D. According to NLM, UMLS is made to provide “health and biomedical vocabularies and standards” to be used and exchanged by computer systems.

Finally, the ontology is publicly available.

From this ontology seven records were randomly picked by running through the ontology records and pick the ones with related relationships. The records were randomly picked from different classes in the ontology view (Appendix C).

IV. Results and Discussion

All the chosen records from the ontology can be considered to be ones of the part_of relation records. We wanted to make this study more comprehensive by not just indicating if the relationship is similar or not, but also measure its similarity. Therefore, if a relationship is accurate (i.e. part_of), its weight is 1, %100 similar. If the relationship can be part_of or located_in, its weight is 0.5, %50 similar. If the relationship should be located_in, the records weight is 0 to indicate it should be the other relationship. Finally, if the relationship cannot be part_of or located_in, the weight is -1. The weight is -1 to indicate it is a crucial error. Please remember that the relationships part_of and located_in are confused in many cases; therefore, it can be considered less of an error compared with not needing to have either of the relations.

The results from the medical dictionary are shown in Appendix B. The dictionary average is calculated as 0.5+(-1)+0+1+(-1)+(-1)+(-1) / 7 = -0.21429 = -%21.429. The -%21.429 is considered the similarity resulted from the medical dictionary. Being from the U.S. National Library of Medicine, National Institutes of Health, and Merriam-Webster, Incorporated we consider the results from MedlinePlus to be the accurate relationship and will be used as the reference. On the other hand, all the records picked from the FMA view of neuroanatomey ontology are considered part_of; see appendix c. Therefore, the average is 1+1+1+1+1+1+1 / 7 = 1 = %100.

The similarity is measured between the results from the medical dictionary (A) and ontology's records (B) by using the similarity measures discussed in section 2.

1) Jaccard = J (A, B) = |A⋂B| / |A⋃B | [23] = 0.666+0+0.333+1+0+0+0 / 14 = 0.14279 = %14.279. In order to consider the weight of the relationships we weighted the relationships as follows:

- The relationship of the value 1 = 1.
- The relationship of the value 0.5 = 0.666.
- The relationship of the value 0 = 0.333.
- The relationship of the value -1 = 0.

2) Ochiai = A.B / $\sqrt{\sum_{i=1}^{n} A_i \cdot \sum_{i=1}^{n} B_i}$ [16] = (1X0.5)+(1X1)+(1X0)+(1X1)+(1X-1)+(1X-1)+(1X-1) / $\sqrt{0.5+1+0+1+1+1+1}$ X $\sqrt{7}$= -2.5 / 6.205 = -0.4029 = -%40.29.

3) For Forbes, Simpson, and Yule measures

- a = Number (A⋃B). The value of a is calculated by adding all readings together, from the medical dictionary and the ontology, while considering the weight the reading. Therefore, a = 0.5+(-1)+0+1+(-1)+(-1)+(-1) + 1+1+1+1+1+1+1 = 4.5
- b = Number (A-B). The value of b is calculated by adding the readings that exist in A, the reading does not equal 0, but does not exist in B. To consider the weight of the readings we calculated as follows:
  - b = 0+(-2)+0+0+(-2)+(-2)+(-2) = -8. The first record is A=0.5 and B=1 which indicates that the whole value of A exists in B so, it should be 0. Now, for the second, fifth, sixth, and seventh records, A=-1 and B=1, we are measuring the existence of -1 in 1. The distance between -1 and 1 is 2. Since we move from -1 we add a negative sign, -2. Finally, for the third (A=0 and B=1) and fourth (A=1 and B=1) records we check the existence in A and the inexistence in B in a straight forward manner
- c = Number (B-A). The value of c is calculated by adding the readings that exist in B, the reading does not equal 0, but does not exist in A. To consider the weight of the readings we calculated as follows:
  - c = 0.5+2+1+0+2+2+2 = 9.5. The first record, A=0.5 and B=1, half of B existence in both A and B and the other half does exit in B but not in A. For the second, fifth, sixth, and seventh records, A=-1 and b=1, the distance between -1 and 1 is 2 and since we are traveling from positive, the result is positive, 2. Finally, the third (A=0 and B=1) and fourth (A=1 and B=1) records are calculated in a straight forward manner
- d = Number (-A-B). The value of d is calculated by adding the number of elements that does not exist in either A or B. Since all elements exist in B, the results is d=0
- n = a+b+c+d. Therefore, n=4.5+(-8)+9.5+0

a) Forbes = na / (a+b) (a+c) [8] = 6X4.5 / [4.5+(-8)](4.5+9.5) = 27 / -49 = -0.55102 = - %55.102

b) Simpson = a / Min(a+b, a+c) [8] = 4.5 / Min(4.5+0, 4.5+9.5) = 4.5 / 4.5 = 1 = %100

c) Yule = ad-bc / ad+bc [8] = 4.5X0 – (-8)X9.5 / 4.5X0 + (-8)9.5 = 76 / -76 = -1 = -%100

V. Concluding Remarks

In this paper, we are contributing by studying biomedical ontology inconsistency using similarity measures. The ontology records were taken from an ontology created by Unified Medical Language System (UMLS). The records picked from the Foundational Model of Anatomy(FMA) view of neuroanatomey ontology. The similarity between the results from the medical dictionary and the ontology records are calculated. In addition, the distance between the results from the measures and the results from the medical dictionary is computed, Table 1.

TABLE I. THE RESULTS FROM THE SIMILARITY MEASURES COMPARED WITH THE RESULTS FROM THE MEDICAL DICTIONARY, ONTOLOGY AVERAGE = -%21.429

Measure	Jaccard	Ochiai	Forbes	Simpson	Yule
Similarity	%14.279	-%40.29	-%55.102	%100	-%100
Difference with the result from the medical dictionary	21.429 + 14.279 = 35.708	40.29 – 21.429 = 18.861	55.102-21.429 = 33.673	100 + 21.429 = 121.429	100 – 21.429 = 78.571

ANOTHER CONTRIBUTION OF THE PAPER IS RANKING SIMILARITY MEASURES THAT ARE MOST USED IN THE BIOMEDICAL FIELD BASED ON THEIR ACCURACY IN MATCHING THE SIMILARITY OF THE ONTOLOGY AVERAGE. BASED ON TABLE 1, THE RANKING OF THE SIMILARITY MEASURES IS SUGGESTED AS FOLLOWS: OCHIAI, FORBES, JACCARD, YULE, AND THEN SIMPSON.

VI. FUTURE WORK

There are four main points I will be doing in the future (in a rough order). First, I will be working on including more relationships and more records in the study. Second, I will find a resolution for ontology inconsistency challenge. Third, I will work on developing a software that can automatically fix the ontologies by apply some of the resolutions like the one suggested by [22] to resolve the challenge of part_of and located_in confusion. Finally, I will increase the number of experts participating in the study.

A. Experts Results

There was not a sufficient number of responses from field experts so the part of using experts is delayed to future publications.

By looking at the projects that use the ontology and the ontology developer in the prospective of the expert's answers, it suggests a major concern of the information included in the ontology. According to the answers, six of the seven sentences (i.e. 1, 3, 4, 5, 6, and 7) can use the relationships part_of or located_in, interchangeably. When keeping in mind that at the end of the day this ontology (and just like any other ontology) is meant to be exchanged by computer systems without (or little) human interference, it indicates that interpretation errors can occur. A computer system may (for example) reject a located_in interpretation since the one that computer system has is part_of. This is a typical scenario of miss communication. If a health insurance company has one or more of the six sentences with the relationship located_in and a hospital (e.g. Johns Hopkins Hospital) uses this ontology and has the relationship as x_part_of, the computer system of one of them (or even both) may not recognize the record sent, not to mention the possibility of miss interpreting the sent record. Finally, according to the expert's answers, sentence number 2 has the wrong relationship and cannot be either part_of or located_in

ACKNOWLEDGMENT *(Heading 5)*

The author thanks Prof. Anthony F. Norcio, Professor of Information Systems, Department of Information Systems, University of Maryland (UMBC) for his support.

REFERENCES

[1] [1] N.A.A. Aziz, S.S. Salleh, D. Mohamad, M. Omar, Investigating Jaccard Distance similarity measurement constriction on handwritten pen-based input digit, in: Science and Social Research (CSSR), 2010 International Conference on, (IEEE, 2010), pp. 1181-1185.

[2] [2] F. Baader, The description logic handbook: theory, implementation, and applications, (Cambridge university press, 2003).

[3] [3] S. Basu, M. Bilenko, R.J. Mooney, Comparing and unifying search-based and similarity-based approaches to semi-supervised clustering, in: Proceedings of the ICML-2003 Workshop on the Continuum from Labeled to Unlabeled Data in Machine Learning and Data Mining, (2003), pp. 42-49.

[4] [4] T. Berners-Lee, J. Hendler, O. Lassila, The semantic web, Scientific american, 284(5) (2001) 28-37.

[5] [5] C. Cassisi, P. Montalto, M. Aliotta, A. Cannata, A. Pulvirenti, Similarity Measures and Dimensionality Reduction Techniques for Time Series Data Mining, Advances in Data Mining Knowledge Discovery and Applications, (2012).

[6] [6] S. Cayzer, Semantic Blogging AND DECENTRALIZED KNOWLEDGE MANAGEMENT, Communications of the ACM, 47(12) (2004) 47-52.

[7] [7] S.-H. Cha, C.C. Tappert, S.N. Srihari, Optimizing binary feature vector similarity measure using genetic algorithm and handwritten character recognition, in: Proceedings of the Seventh International Conference on Document Analysis and Recognition, (Citeseer, 2003), pp. 662.

[8] [8] S.-S. Choi, S.-H. Cha, C. Tappert, A survey of binary similarity and distance measures, Journal of Systemics, Cybernetics and Informatics, 8(1) (2010) 43-48.

[9] [9] T. Clifford, W. Stephennson, An introduction to numerical classification, (Academic Press, 1975).

[10] [10] J. Domingue, D. Fensel, J.A. Hendler, Handbook of semantic web technologies, (Springer, 2011).

[11] [11] B. Fei, Z. Lee, J.L. Duerk, D.L. Wilson, Image registration for interventional MRI guided procedures: Interpolation methods, similarity measurements, and applications to the prostate, in: Biomedical Image Registration, (Springer, 2003), pp. 321-329.

[12] [12] S. Forbes, Method of determining and measuring the associative relations of species, Science, 61(1585) (1925) 518-524.

[13] [13] S.A. Forbes, On the local distribution of certain Illinois fishes: an essay in statistical ecology, (Illinois State Laboratory of Natural History, 1907).

[14] [14] V.R.B.a.P.C.a.J.B.a.A. Gangemi, Law and the Semantic Web, an Introduction, (2003).

[15] [15] P. Jaccard, Etude comparative de la distribution florale dans une portion des Alpes et du Jura, (Impr. Corbaz, 1901).

[16] [16] D.A. Jackson, K.M. Somers, H.H. Harvey, Similarity coefficients: measures of co-occurrence and association or simply measures of occurrence?, American Naturalist, (1989) 436-453.

[17] [17] R. Jongman, C. Ter Braak, O. Van Tongeren, R. 1995. Data analysis in community and landscape ecology, in, (Cambridge Univ. Press).

[18] [18] A. Lozano-Tello, A. Gómez-Pérez, Ontometric: A method to choose the appropriate ontology, Journal of Database Management, 2(15) (2004) 1-18.

[19] [19] V. Maojo, J. Crespo, M. García-Remesal, D. De la Iglesia, D. Perez-Rey, C. Kulikowski, Biomedical ontologies: toward scientific debate, Methods of information in medicine, 50(3) (2011) 203.

[20] [20] V. Maojo, M. García-Remesal, H. Billhardt, R. Alonso-Calvo, D. Pérez-Rey, F. Martín-Sánchez, Designing new methodologies for integrating biomedical information in clinical trials, Methods of information in medicine, 45(2) (2006) 180.

[21] [21] A. Ochiai, Zoogeographic studies on the soleoid fishes found in Japan and its neighbouring region, Bull. Jpn. Soc. Sci. Fish. (Nihon Suisan Gakkaishi), (22) (1957) 526-530.

[22] [22] O. Rabie, A.F. Norcio, Discussion of Some Challenges Concerning Biomedical Ontologies, in: Human-Computer Interaction. Applications and Services, (Springer, 2013), pp. 173-180.

[23] [23] D. SáNchez, M. Batet, Semantic similarity estimation in the biomedical domain: An ontology-based information-theoretic perspective, Journal of biomedical informatics, 44(5) (2011) 749-759.

[24] [24] S. Schulz, A. Kumar, T. Bittner, Biomedical ontologies: What part-of is and isnÕt, Journal of Biomedical Informatics, 39(2006) 350-361.

[25] [25] G.G. Simpson, Notes on the measurement of faunal resemblance, American Journal of Science, 258(2) (1960) 300-311.

[26] [26] V. Sugumaran, V.C. Storey, Ontologies for conceptual modeling: their creation, use, and management, Data & Knowledge Engineering, 42(3) (2002) 251.

[27] [27] V.V. Vavilova, M. William Jr, Temporal and altitudinal variations in the attached algae of mountain streams in Colorado, Hydrobiologia, 390(1-3) (1998) 99-106.

[28] [28] P. Willett, Similarity-based approaches to virtual screening, Biochemical Society Transactions, 31(2003) 603-606.

[29] [29] P. Willett, Structural biology in drug metabolism and drug discovery, Biochemical Society Transactions, 31(part 3) (2003).

[30] G.U. YULE, Notes on the theory of association of attributes in statistics, Biometrika, 2(2) (1903) 121-134.

Appendix A: Experts results

The following is the exact text that was sent to experts as an attached file via email:

"The following are seven sentences that contain biomedical terms. In each sentence there are two biomedical terms with a blank (i.e. …) in between.
The study is to test the relationships "part_of" and "located_in". For each sentence pick one of the options "part_of" (i.e. the blank should be filled with "part_of"), "located_in" (i.e. the blank should be filled with "located_in"), either (i.e. the blank should be filled with either "part_of" or "located_in"), or neither (i.e. the blank should be filled with neither "part_of" nor "located_in").
1- Hypoglossal Never Tract … Medulla Oblongata.
2- Supracallosal Gyrus … Limbic Lobe.
3- Left Intermediate Gray Matter … Gray Matter of Spinal Cord.
4- Midbrain … Brain.
5- Hypoglossal Nerve Tract … Medulla Oblongata.
6- Dentate Gyrus of Right Hippocampus … Right Hippocampal Formation.
7- Terminal Ventricle … Central Canal of Spinal Cord."

The expert is an Infectious Disease Specialist at King Abdulaziz Medical City-National Guard Hospital, Jeddah, Kingdom of Saudi Arabia. The expert's answers are as follows:

"1- Hypoglossal Never Tract … Medulla Oblongata. (either)
2- Supracallosal Gyrus … Limbic encephalitis. (neither)
3- Left Intermediate Gray Matter … Gray Matter of Spinal Cord. (either)
4- Midbrain … Brain. (either)
5- Hypoglossal Nerve Tract … Medulla Oblongata. (either)
6- Dentate Gyrus of Right Hippocampus … Right Hippocampal Formation. (either)
7- Terminal Ventricle … Central Canal of Spinal Cord. (either)."

Appendix B: Seven

Number seven has a unique indication in the Arabian culture. The numbers seven and seventy implies completeness, plentifulness, or enough. A suggested reason is that number seven consists of the addition of an even number (i.e. 4) and an odd number (i.e. 3) without repetition. Many translate the number seventy literally without its semantic implication. If the phrase "you are a chicken" translated literally into another language, the recipient of the translation may be confused for calling him a chicken (i.e. a bird). He may understand the phrase intention better if it was translated semantically (i.e. to become "you are a coward").

Appendix C: The Medical Dictionary

The following are the results from MedlinePlus by the U.S. National Library of Medicine, National Institutes of Health, and Merriam-Webster, Incorporated:

1- hypoglossal nerve: either of the 12th and final pair of cranial nerves which are motor nerves arising from the medulla oblongata and supplying muscles of the tongue and hyoid apparatus in higher vertebrates—called also hypoglossal, twelfth cranial nerve.

2- Limbic Lobe: the marginal medial portion of the cortex of a cerebral hemisphere.

3- Gray Matter: neural tissue especially of the brain and spinal cord that contains cell bodies as well as nerve fibers, has a brownish gray color, and forms most of the cortex and nuclei of the brain, the columns of the spinal cord, and the bodies of ganglia—called also gray substance.

4- Midbrain: the middle division of the three primary divisions of the developing vertebrate brain or the corresponding part of the adult brain that includes a ventral part containing the cerebral peduncles and a dorsal tectum containing the corpora quadrigemina and that surrounds the aqueduct of Sylvius connecting the third and fourth ventricles—called also mesencephalon.

5- Substantia Nigra: a layer of deeply pigmented gray matter situated in the midbrain and containing the cell bodies of a tract of dopamine-producing nerve cells whose secretion tends to be deficient in Parkinson's disease.

6- Dentate Gyrus: a narrow strip of cortex associated with the hippocampal sulcus that continues forward to the uncus.

7- Central Canal: a minute canal running through the gray matter of the whole length of the spinal cord and continuous anteriorly with the ventricles of the brain.

Appendix D: Ontology Records

The records picked from the FMA view of neuroanatomey ontology:

1- Hypoglossal Nerve Tract constitutional_part_of Medulla Oblongata.

2- Supracallosal Gyrus regional_part_of Limbic Lobe.

3- Left Intermediate Gray Matter regional_part_of Gray Matter of Spinal Cord.

4- Midbrain regional_part_of Brain.

5- Right Substantia Nigra Pars Compacta regional_part_of Right Substantia Nigra.

6- Dentate Gyrus of Right Hippocampus regional_part_of Right Hippocampal Formation.

7- Terminal Ventricle regional_part_of Central Canal of Spinal Cord.

Biography

Osama Bassam J. Rabie is currently a fulltime Ph.D. in Business with a concentration in Information Systems student at School of Business at Virginia Commonwealth University. Mr. Rabie is a teacher assitant at Department of Information Systems, Faculty of computers and information technology, King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia. He holds a Bachelor of Science degree in Computer Science from King Abdulaziz University and a Master of Science degree in Information Systems from University of Maryland (UMBC).