Domain Independent Integrated Multimedia Data Mining is a Frivolous Exercise: Classicism vs. Connectionism Debate Revisited

Muhammad A Ahmad and Ankur M Teredesai

Department of Computer Science
102 Lomb Memorial Drive,
Rochester Institute of Technology,
Rochester, NY 14623
{maa2454, amt}@cs.rit.edu

Abstract. In this paper the classicist and the connectionist debate is integrated in a multimedia data mining framework. Due to the ready availability of visual information digitally, the problem to store, retrieve and organize images is gaining tremendous importance and poses a serious challenge for practitioners of AI and databases alike. The central problem for image database management is the idea of an annotation associated with every image. This idea of an annotation as a representation of an image is explored in this paper. The claim is that arriving at a suitable annotation using data mining techniques however does not solve the problem of representation and reference. Most image annotation techniques are classicist, while connectionist approaches which have the advantage of contextuality have not been explored much because of the problem of systematicity. Consequently the image annotation problem can be used as a critique of both.

1 Introduction

Knowledge Representation (KR henceforth) deals with the problem of representing and storing knowledge. Hence, KR has a strong affinity with Epistemology. KR has been one of the foremost problems in AI since its inception. A representation is a surrogate for the thing itself [7]. Consequently, the way information is represented also effects how it is processed and this is why KR is one of the main problems of AI. In this paper we consider these problems from a data mining perspective and explore what AI may learn from this perspective.

McCarthy et. al [15] considered the KR problem in AI from a philosophical point of view. The important questions that they considered can be characterized as follows: finding optimal representations of the (physical) world that accommodates addition of new laws or facts, representations for non-physical entities, knowledge acquisition and representing this knowledge to oneself. It can be reasonably stated that Data Mining has inherited most of these problems from AI and although Data Mining is a derivative field of AI, they do have subtle differences. Data Mining puts its emphasis on knowledge discovery in massive amounts of data, whereas AI emphasizes learning and reasoning the unknown. Consequently in Data Mining KR is tied to usability. The task of representing the world thus becomes the task of finding representations that can reveal non-obvious knowledge about the world, knowledge acquisition becomes filtering knowledge from an already available corpus of knowledge. Finally the importance of presenting knowledge to the self is debatable since a human may or may not be present in the information processing loop.

Scalability is perhaps the most severe problem for any data mining technique [4]. The problem of scalability is rooted in computational complexity [17], limitations in KR and the non-generalizability of search techniques. Also, the role of domain knowledge to reduce the search space for mining algorithms has been considered before [2] [19]. Since algorithmic complexity of general data mining techniques is beyond the scope of this discussion we shall consider the other two factors and their relation to one another.

The two dominant paradigms in Knowledge Representation are the connectionist and the classicist paradigms. Most data mining algorithms are classicist in the sense that knowledge is almost always represented as symbols or in computational terms as bits of strings. This is especially pronounced in scenarios where the data is well structured, although the "solutions" are sometimes marred by scalability problems. The problem recurs in connectionist systems also. In selecting an architecture for solving a particular mining problem, the scalability of its implementation in the particular context is generally taken into consideration. It would not be too much of an exaggeration to state that the classicist vs. the connectionist debate is really between local vs. distributed representations.

Connectionism primarily draws its inspiration from the human brain, although associationism which goes as far back as David Hume subsumes connectionism. Connectionism claims that knowledge is represented in a distributed non-symbolic manner in the connection strengths between neurons. Connectionist architectures like neural networks have been employed for data mining since its beginning, although the structured nature of many mining problem makes it appear as if they are not susceptible to connectionist solutions. There is a middle ground between the two paradigms, the so called implementational connectionism for whom the the aim of connectionism is to find connectionist architectures that implement classicist architectures at a higher more abstract level. This line of research has not been pursued much in data mining, where it can be applied to problems like auto annotation of images. Perhaps the greatest drawback that connectionism has is its lack of systematicity [11], although a line of reasoning [6] [5] suggests that the problem can be overcome by implementational connectionism.

2 Annotation as a Representation

The task of image annotation is trivial for a human being but when the task is transferred to the machine learning domain it becomes challenging [18] [16]. Traditionally, image mining techniques involving higher-level object detection and complex feature extraction have been employed for scene understanding, though what constitutes the semantic content of understanding can be disputed. Text descriptions for a given image are then generated based on this "understanding." An annotation can thus be thought of as a representation, a surrogate for the image itself. An annotation represents either objects present in the image or a situation. Hence annotation should not be however confused with naming objects, although the former subsumes the latter. This falls nicely into the classicist paradigm *i.e.,*manipulation of symbols and rules to give suitable symbols as answers to queries which itself is a set of symbols.

We have reframed the problem of image auto-annotation to that of multi-relational association rule mining where relations exist between image-based features and textual annotations. We then combined low-level image features such as color, orientation, intensity, etc. and corresponding text annotations to generate multi-relational association rules that were used to generate annotations. The relations between these media are not rigid. The idea is that annotations expressed as multiple relations stores more information as compared to when these are considered in isolation.

3 The CoMMA Framework

In a previous work [18] we developed a framework, CoMMA, for automating the task of image annotation by using multi-relational mining. The MR-FP tree algorithm [18] was devised for finding image annotations. The MR-FP Tree algorithm is the multi-relational version of the FP Tree algorithm [13] which is an association rule mining algorithm. We restrict the images under consideration to certain domains, hence it shows some affinity to the domain specificness of expert systems. The low-level image features that were considered for mining are color (red, green, blue and yellow), orientation (edge orientations of $0°$, $45°$, $90°$ and $135°$) and intensity. The image features are extracted based on Itti and Koch's [14] focus of attention theory. The selective attention model allows the system to concentrate on processing salient objects in the scene without the need to process unimportant aspects. Input images are processed in three parallel feature channels by the attention model processes. The feature saliency maps topographically represent the saliency of objects in the scene based on respective features. For a detailed description of still-feature extraction from images, refer to Gaborski's previous work [12] on the subject.

The MRFP-Growth algorithm that we have used in CoMMA, builds upon the FP-Growth algorithm and consists of two phases. The first phase of the algorithm involves running the MR FP-Tree algorithm separately on all the tables. This phase is similar to the FP Tree algorithm but with one major difference. Each node in the tree not only keeps track of its support but also keeps track of the indices in the dataset where it occurs. In the second phase the table that was populated with rules mined from MR FP-Trees in the previous phase is used to make an MR FP-tree. Once the tree is made, it is mined for rules as specified earlier. The task of MR FP-Tree can thus be stated as that of finding relations from the primary table to each of the secondary tables and then between the secondary tables. Refer to our CoMMA paper [18] for a detailed discussion of the MR-FP Tree algorithm. One of the inspirations for CoMMA was connectionism.

Original Keywords: small bush flowers ground
Auto-Annotations: small bush grass leaves radiant flowers berries ground weeds earth dry desert

Original Keywords: bush flowers ground
Auto-Annotations: small bush over forest river ground tree leaves bushes grass sky flowers trees rocks

Original Keywords: lake summit in Alaska
Auto-Annotations: still lake summit Alaska near rock rocks desert melting ice walls

Fig. 1. Some Images and their Corresponding Annotations

The relations between images, their annotations and the image features are more than just statistical correlations and the relations are not rigidly defined. The system overall fits into the traditional classicist scheme. The results of some of the CoMMA generated annotations are given in Figure 1. From these annotations it is evident that the system picks up major objects present in the picture, though the annotations are not really perfect. The results vary as the support and confidence is varied. Hence the task is to get an optimal value for these two parameters for a particular database.

4 Turing Test Revisited

Ever since Alan Turing formulated his test for intelligence other versions of the test have been proposed like the Total Turing Test, the Reverse Turing Test, the modified Turing Test, the Limited or constrained Turing Test. One can reframe the problem of image annotation to that of a constrained Turing Test where the task is to find appropriate words to describe an image. Lets call it the "Annotation Turing Test." Although a system that can annotate images in an open ended domain is currently impossible, the Annotation Turing Test can be conducted in a limited domain, say dealing with only scenic images or sports images *i.e.,*a Constrained Annotation Turing Test (CATT henceforth). Section 3 described a system for automatically annotating images and although the system is far from perfect in auto-annotation of images but the fact that it and other similar systems can annotate images with reasonable accuracy gives one reasons to believe that a domain specific auto-annotation system might not be that far off.

Now consider the question, is the system said to have the ability to understand images if it can properly annotate them? Admittedly the test has some limitations but since CATT is a constrained test to begin with, this is a non-issue. If one takes a purely functionalist stance then one is forced to say that it does, but there is clearly something wrong with this assertion. The problem arises because of the confusion between cognition and representation. The Problem with CATT is that seemingly intelligent behavior exhibited by the system is not generative. The system is brittle if moved from one domain to the next. The Turing Test would be vindicated if a general purpose non-domain specific system could be built but the very notion of such a system has many problems. The reason that humans have a need to name objects (another way of saying annotating) is because they are situated in an environment where they have to interact with the world and that these objects have some use for them, in short a human is, as Heidegger puts it a being-in-the-world or dasein. This is perhaps the reason why the Holy Grail of AI *i.e.,* general purpose reasoning system has been so elusive.

5 Data Mining: Classicism or Connectionism?

Data mining seeks to extract hidden knowledge from large amounts of data. But what constitutes hidden knowledge really? Most traditional data mining algorithms are either search algorithms or their derivatives. Hence data mining is confronted with the same problems that earlier plagued many classicist architectures *i.e.,* the problem of determining what is relevant. In data mining tweaking relevant parameters in an algorithm usually ensures that one gets the desired results *e.g.,* consider the case of Multi-relational Association Rule Mining described in section [18], changing support and confidence affects the results given by the image annotation system. This points toward a connectionist solution *i.e.,*optimum values for these and other parameters might be found by a connectionist system, one that would work in multiple domains though not always. Such a hybrid solution would also give us reasons to believe that image annotation is not merely a search task.

Classicists architectures also fail to account for the role of context in real world problems. Connectionism offers the hope of doing away with the problem of holistic representations [6], although it could be that systematicity might be the trade off in this case. The systematicity problem is an outstanding problem for connectionism for which a solution has not yet been devised. In classicism, a solution for systematicity comes free as it can be easily made part of the architecture itself. Hence the scope of the problem determines the architectures of the solution. Data Mining posits that there are non-obvious hidden patterns in data, hence there are multiple levels of interpretation and complexity within the data. The type of patterns that are discovered are also dependent upon the type of representations that are used; changing KR changes the patterns discovered.

6 Unsolved Problems

What AI can learn from data mining is that although domain specific human capabilities can be mimicked by search methods, these are just search mechanisms and nothing more. This is of course the lesson which should have been learnt from the Expert System epoch in AI. It can be interpreted as either a short coming of functionalism or a failure to distinguish between cognition and understanding. We have stated the shortcomings of the classicist and the connectionist paradigms with respect to the problem of image annotation and what generalizations can we make from the lessons learned here. The symbol manipulation paradigm of Classicism has also been attacked vigorously by Dreyfus [10]. Dreyfus' critique rests on the premise that human intelligence is embodied. Dreyfus elaborates this idea by noting that "Merleau-Ponty's notion of intentional arc is meant to cover all three ways our embodied skills determine the way things show up for us. [9]" Humans recognize objects and situations which they are confronted with, image annotation is about the reference to these. Hence, situatedness may be the hallmark of intelligence.

Dreyfus has pointed out that classicism is in part motivated by the rationalist tradition in Philosophy. Hence an image annotation system can annotate images with a reasonable accuracy but such a system refers objects and situations in the image in the same sense that the images refer to these objects in the physical world. Thus the problem is that of intentionality. Dreyfus however thinks [9] [8] that the way in which connectionist systems learn shows some similarity to how the intentional arc is established in humans. If correct then the assertion hints how the problem of reference might be solved. It is interesting to note that although Dreyfus favors the connectionist approach, he has cast doubt by asserting [8] [9] that even this might not be sufficient to do the job. While one can substitute real world problems with toy problems or even "solve" these problems within a restricted domain, a general problem solver even within a specific domain has remained elusive. It could be that to seriously tackle this problem we would have to eventually abandon the dichotomy of disjoint abilities for accounting human intelligence. Solving the problem of contextuality and specificity, a common problem in image annotation, simultaneously might require that we augment the search mechanism with other capabilities. Machine learning and computational learning theorists are exploring ensembles of search mechanisms to provide associative information in order to enhance the search capabilities of traditional algorithms [3] [1]. However, the utility of these capabilities remains to be seen.

As argued in section 4 a human is a being-in-the-world. Her intelligence is in part due to her interaction with the world. This is why an open-ended, non domain specific annotation system is a too difficult and complex problem to be solved by any of the currently available paradigms. Of course it would not be practical to build a system that learns annotations from the world by physically interacting with it. Such a project would of course the defeat the purpose of auto annotation in the first place. current techniques in image annotation like association rule mining, object identification, viewpoints etc. perform reasonably in this task. Algorithmic improvements on these approaches will certainly yield better results, even though the domain restrictions will imply nonetheless.

7 Conclusion

Many of the problems that are traditionally associated with AI were considered from a data mining perspective by specifically considering the image mining problem. The strengths and weaknesses of the classicist and the connectionist paradigms were considered. The contextualization problem in classicism is a major impedent to generalization. This is one of the main reasons why image annotation systems have to restricted to certain image domains. Connectionism works best in situations where the KR is distributed. One has to consider the trade offs when considering the problem to be solved.

An image annotation can also be thought of as a representation of the image itself. We also argued that system that annotates images cannot be said to refer to anything in the image, where reference is an intentional construct. CATT also demonstrates that a genuinely referring system has to understand what it is referring. An open ended image annotation system can thus on be part of a larger general problem solving system. It does not have to be similar to Heidegger's dasein or being-in-the-world. However the establishment of Merleau Ponty's intentional arc might be a necessary condition. While considering the problem of image annotation we consider our own newly devised system CoMMA. CoMMA uses multi-relational association rules to come up with annotations for images. The system performs reasonably well within a restricted domain. However the performance deteoriates as more domains are added. Although CoMMA is classicist, the relations between annotations can be thought of as associationalist if not connectionist.

What does the Turing test really tell us about the intelligence of a system or even about intelligence at all? A debate on CATT revealed that particular intelligent capabilities are not isolated but are always part of the whole. Since a human, a being-in-the-world according to Heidegger, is always situated in an environment. It is due to her interaction with the world that meaning and reference arises, including naming objects and situations. Hence situatedness is be a necessary condition for any general level intelligent capability like annotating in an open ended domain.

Possible arenas of research for the future include extending CoMMA to give descriptions of the images, test the performance of the system in multiple disparate domains and improve upon the MR FP Tree algorithm. Perhaps the long term goal of such a project could be to test an implementational connectionist architecture or some other hybrid connectionist-classicist system that can give us a better idea about what constitutes a reference. The relation between reference and the establishment of an intentional arc should be explored.

References

1. K Ali *A comparison of methods for learning and combining evidence from multiple models*University of California, Irvine, Dept. of Information and Computer Sciences.
2. Sarbjot Anand, David A. Bell, John G. Hughes *The Role of Domain Knowledge in Data Mining*, CIKM, Baltimore, Maryland, 1995.
3. Martin Anthony *Generalisation Error Bounds for Threshold Decision Lists* Journal of Machine Learning Research, 2004, 189–217
4. Hendrik Blockeel, Michle Sebag *Scalability and efficiency in multi-relational data mining* In ACM SIGKDD, Special Issue on Multi-Relational Data Mining, Vol 5, p. 17-30.
5. David Chalmers *Syntactic transformations on distributed representations*http://citeseer.ist.psu.edu/chalmers90syntactic.html 1990
6. L Chrisman *Learning Recursive Distributed Representations for Holistic Representation*, Connection Science vol.3, no 4., 1991, 345 – 366, Carfax Publishers.
7. Randall Davis, Howard Shrobe, and Peter Szolovits *What is A Knowledge Representation?* AI Magazine 14 (1), 1993, 17–33
8. Hubert L Dreyfus, *Merleau-Ponty's Critique of Mental Representation: The Relevance of Phenomenology to Scientific Explanation Intelligence Without Representation* Houston Studies in Cognitive Science. http://www.hfac.uh.edu/cogsci/dreyfus.html 1998
9. Hubert L Dreyfus, *The Current Relevance of Merleau-Ponty's Phenomenology of Embodiment* The Electronic Journal of Analytic Philosophy http://www.phil.indiana.edu/ejap/1996.spring/dreyfus.1996.spring.html 1996
10. Hubert L Dreyfus, *What computers still can't do: a critique of artificial reason*, MIT Press, Cambridge, MA, 1992
11. J Fodor and Z Pylyshyn, *Connectionism and Cognitive Architecture: a Critical Analysis*, Cognition, 28 1988: 3–71
12. R Gaborski, V S Vaingankar, and R L Canosa, *Goal Directed Visual Search Based on Color Cues: Cooperative Effects of Top-down and Bottom-up Visual Attention.* In Proceedings of the Artificial Neural Networks in Engineering, Rolla, Missouri. Vol 13, 613–618, 2003
13. J Han, J Pei, and Y Yin *Mining frequent patterns without candidate generation* In ACM-SIGMOD, Dallas, 2000.
14. L Itti , and C Koch *Computational modeling of visual attention.* Nature Neuroscience Review., 2(3):194-203. 2001
15. J McCarthy and P Hayes. *Some philosophical problems from the standpoint of artificial intelligence.* In Meltzer and Michie, editors, Machine Intelligence, 4, 463–502. Edinburgh University Press, 1969.
16. Jia-Yu Pan, Hyung-Jeong Yang, Christos Faloutsos and Pinar Duygulu *Automatic Multimedia Cross-modal Correlation Discovery* KDD 2004, Seattle, WA, Aug. 2004.
17. B Selman and H Kautz. *Knowledge compilation and theory approximation* Journal of the ACM, 43(2):193–224, Mar. 1996. http://citeseer.ist.psu.edu/article/selman96knowledge.html
18. Ankur Teredesai, Muhammad Ahmad, Juveria Kanodia, and Roger Gaborski. *COMMA: Combing Multimedia Multi-relational Associations* In The 7th ACM SIGKDD Workshop on Multimedia Data Mining, Seattle, WA, USA 2004.
19. Suk-Chung Yoon, Lawrence J Henschen, E K Park, S A M. Makki *Using Domain Knowledge in Knowledge Discovery*CIKM 1999: 243-250