Towards Bisociative Knowledge Discovery⋆

Michael R. Berthold

Nycomed Chair for Bioinformatics and Information Mining,
Department of Computer and Information Science,
University of Konstanz, Germany
Michael.Berthold@Uni-Konstanz.DE

Abstract. Knowledge discovery generally focuses on finding patterns within a reasonably well connected domain of interest. In this article we outline a framework for the discovery of new connections between domains (so called *bisociations*), supporting the creative discovery process in a more powerful way. We motivate this approach, show the difference to classical data analysis and conclude by describing a number of different types of domain-crossing connections.

1 Motivation

Modern knowledge discovery methods enable users to discover complex patterns of various types in large information repositories. Together with some of the data mining schema, such as CRISP-DM and SEMMA, the user participates in a cycle of data preparation, model selection, training, and knowledge inspection. Many variations on this theme have emerged in the past, such as Explorative Data Mining and Visual Analytics to name just two, however the underlying assumption has always been that the data to which the methods are applied to originates from one (often rather complex) domain. Note that by *domain* we do not want to indicate a single feature space but instead we use this term to emphasize the fact that the data under analysis represents objects that are all regarded as representing properties under one more or less specific aspect. Multi View Learning [19] or Parallel Universes [24] are two prominent types of learning paradigms that operate on several spaces at the same time but still operate within one domain.

Even though learning in multiple feature spaces (or views) has recently gained attention, methods that support the discovery of connections across previously unconnected (or only loosely coupled) domains have not received much attention in the past. However, methods to detect these types of connections promise tremendous potential for the support of the discovery of new insights. Research on (computational) creativity strongly suggests that this type of out-of-the-box thinking is an important part of the human ability to be truly creative. Discoveries such as Archimedes' connection between weight and (water) displacement and the – more recent – accidental ("serendipitous") discovery of Viagra are two illustrative examples of such domain-crossing creative processes.

⋆ Extended version of [1].

In this introductory chapter we summarise some recent work focusing on establishing a framework supporting the discovery of domain-crossing connections continuing earlier work [3]. In order to highlight the contrast of finding patterns within a domain (usually associations of some type) with finding relations across domains, we refer to the term *bisociation*, first coined by Arthur Koestler in [13]. We argue that *Bisociative Knowledge Discovery* represents an important challenge in the quest to build truly creative discovery support systems. Finding predefined patterns in large data repositories will always remain an important aspect, but these methods will increasingly only scratch the surface of the hidden knowledge. Systems that trigger new ideas and help to uncover new insights will enable the support of much deeper discoveries.

2   Bisociation

Defining bisociation formally is, of course, a challenge. An extensive overview of related work, links to computational creativity and related areas in AI, as well as a more thorough formalisation can be found in [7]. Here we will concentrate on the motivational parts and only intuitively introduce the necessary background.

Boden [4] distinguishes among three different types of creative discoveries: Combinatorial, Exploratory, and Transformational Creativity. Where the second and third category can be mapped on (explorative) data analysis or at least the discovery process within a given domain, Combinatorial Creativity nicely represents what we are interested in here: the combination of different domains and the creative discovery stemming from new connections between those domains.

Informally, bisociation can be defined as (sets of) concepts that bridge two otherwise not –or only very sparsely– connected domains whereas an association bridges concepts within a given domain. Of course, not all bisociation candidates are equally interesting and in analogy to how Boden assesses the interestingness of a creative idea as being new, surprising, and valuable [4], a similar measure for interestingness can be specified when the underlying set of domains and their concepts are known. Going back to Koestler we can summarise this setup as follows:

> *"The creative act is not an act of creation in the sense of the Old Testament. It does not create something out of nothing; it uncovers, selects, re-shuffles, combines, synthesises already existing facts, ideas, faculties, skills. The more familiar the parts, the more striking the new whole."*

Transferred to the data analysis scenario, this puts the emphasis on finding patterns across domains whereas finding patterns in the individual domains themselves is a problem that has been tackled already for quite some time. Put differently, he distinguishes associations that work within a given domain (called *matrix* by Koestler) and are limited to repetitiveness (here: finding other/new occurrences of already identified patterns) and bisociations representing novel connections crossing independent domains (matrices).
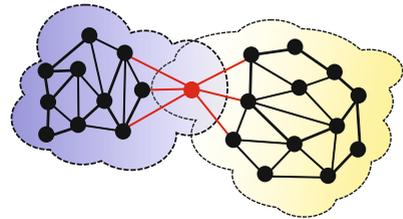
3   Types of Bisociation

Obviously the above still remains relatively vague and for concrete implementations the type of bisociative patterns that are sought needs to be specified better. In the past years a number of bisociation types emerged in the context of Bisociative Knowledge Discovery: Bridging Concepts, Bridging Graphs, and Bridging by Structural Similarity, see [14] for a more detailed analysis. Since these ideas are also addressed in other areas of research, additional types most likely exist in those fields as well.
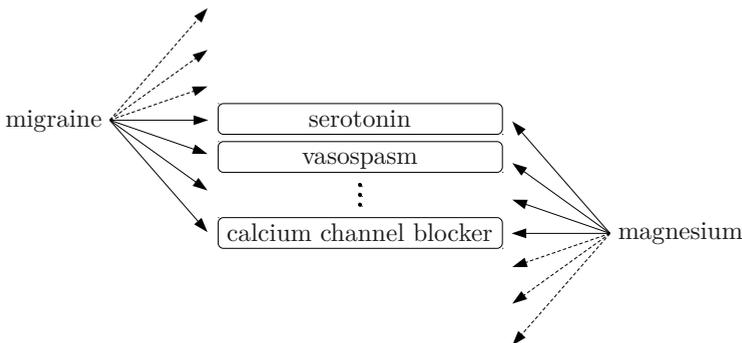
3.1   Bridging Concepts

The most natural type of bisociation is represented by a concept linking two domains, Figure 1 illustrates this.

Such bridging concepts do not need to exist in the context of a network based representation, as suggested by the figure, but can also be found in other representations. In [21], for instance, different textual domains were analysed to find bisociative terms that link different concepts from the two domains.

Fig. 1. Bridging concept (from [14])

An example of a few bridging concepts is shown in Figure 2. Here a well known data set containing articles from two domains (migraine and magnesium) was searched for bridging terms (see [21] for more details). Note that this example reproduces an actual discovery in medicine.

Fig. 2. Bridging concepts - an example reproducing the Swanson discovery (from [21])
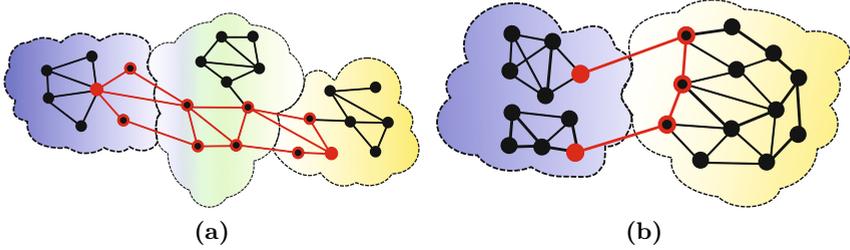
Fig. 3. Bridging graphs (from [14])

3.2   Bridging Graphs

More complex bisociations can be modelled by bridging graphs, Figure 3 illustrates this concept in a network context.

Here two different domains are connected by a (usually small) subset of concepts that have some relationship among themselves. In a network-based representation, a relatively dense subgraph can be identified connecting two domains. However, also in other representations, such "chains of evidence" can be formalised, connecting seperate domains.

Two examples for bridging graphs are shown in Figure 4 (the data stems from Schools-Wikipedia, see [17] for details). These demonstrate well how the two concepts "probability space" and "arithmetic mean" connect the domain of movies with a number of more detailed concepts in the statistics domain. This is at first glance surprising but finds its explanation in the (in both cases also somewhat "creative") use of those concepts in the two films or the series of films dominated by one actor. The second example nicely bridges physical properties and usage scenarios of phonographs.
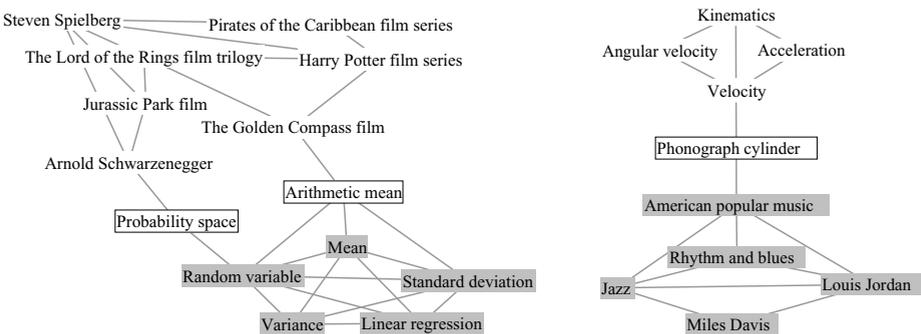

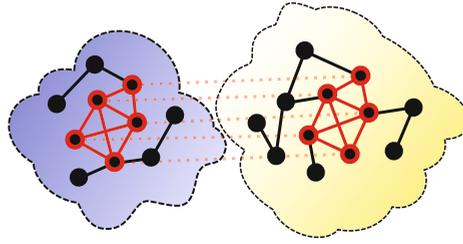**Fig. 4.** Bridging graphs - two examples (from [17])

Fig. 5. Bridging by graph similarity (from [14])

3.3 Bridging by Structural Similarity

The third, so far most complex type of bisociation does not rely on some straight-forward type of link connecting two domains but models such connections on a higher level. In both domains two subsets of concepts can be identified that share a structural similarity. Figure 5 illustrates this – again in a network-based representation; also here other types of structural similarity can exist.

An interesting example of such structural similarities can be seen in Figure 6. The demonstration data set based on Schools-Wikipedia was used in this example again. The two nodes slightly off centre ("Euclid" on the left and "Plato" on the right) are farther apart in the original network but share structural properties such as being closely connected to the hub of a subnetwork ("mathematics" vs. "philosophy"). Note that "Aristotle" also fills a similar role in the philosophy domain.
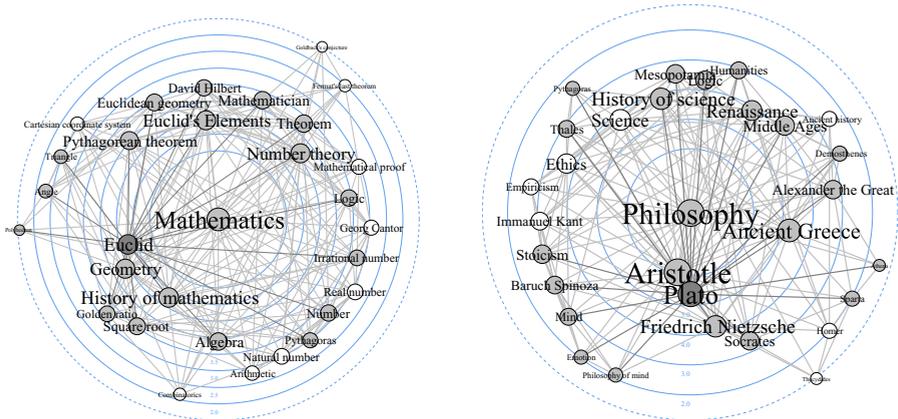

Fig. 6. Bridging by graph similarity - example (from [22])

3.4   Other Types of Bisociation

The bisociation types discussed above are obviously not complete. The first two types are limited to a 1:1 match on the underlying structures and require that the two domains already have some type of (although sparse) neighbourhood relation. Only the third type allows matching on a more abstract level, finding areas of structural similarity and drawing connections between those. Other, more abstract, types of bisociation certainly exist but also more direct bisociation types can be defined as well. This is an exciting area of research and one could also imagine systems that observe user interaction and learn new complex types of bisociation from user feedback and successful discoveries.

4   Bisociation Discovery Methods

In order to formalise the types of bisociations and develop methods for finding them, a more detailed model of the knowledge space needs to be available. When dealing with various types of information and the desire to find patterns in those information repositories a network-based model is often an appropriate choice due to its inherent flexibility. A number of methods can be found in this volume [2]. We hasten to add, however, that this is not the only way to model domains and bisociations, some contributions finding bisociation in non-network type domains can be found here as well, see for example the text-based bisociative term discoveries in [12,21].

It is interesting to note that quite a few of the existing methods in the machine learning and data analysis areas can be used, frequently with only minor modifications. For instance, methods for item set mining can be applied to the detection of concept graphs [15] and measures of bisociation strength can also be derived from other approaches to model interestingness [20,22]. Bisociative Knowledge Discovery can rely to a fairly large extent on existing methods, however the way in which these methods are applied is often radically different. Instead of searching for patterns that have reasonably high occurrence frequencies we are often interested in the exact opposite: the essenace of bisociations is something that is new and whose existence is only hinted at, if at all so far.

This focus on "finding the unexpected" obviously also requires rather different approaches to the creation, analysis and exploration of the underlying structure. Overviews of these three aspects can be found in [5], [23], and [18,9] respectively. Note that an even bigger challenge as opposed to usual knowledge discovery setups is the lack of comprehensive benchmarks. Finding the unexpected is a moving target – once knowledge becomes common sense, it ceases to be all that surprising. In [16] a number of application examples and attempts at benchmarking are summarised and yet there is still scope for work here, specifying how such discovery support systems can be evaluated more comprehensively. The classic setup of benchmark repositories is unlikely to be sufficient, as pure numerical performance does not really quantify a method's potential for creativity support – in fact individual methods will be hard to be evaluated properly, as they only become useful in concert with a larger system enabling truly explorative use.

5   Future Directions

The work briefly outlined in this paper is only the start, of course. Much more needs to be done to fully understand and make use of bisociative knowledge discovery systems. For once, the nature of bisociative patterns is far from complete – so far we have mainly addressed more classical approaches of finding numerical ways to assess the potential for bisociative insights of fairly simple patterns. The true potential lies in finding more abstract techniques to discover bisociations, similar to the methods described in [22] or [11]. Using abstract features to describe neighbourhoods – quite similar to the fingerprint similarity measures used in molecular searches for a long time already – shows enormous promise. Finding structurally similar patterns in different domains allows more complex knowledge to be transferred among the involved domains than only pointing to an existing (or missing) link.

However, in order to support the exploration of these more complex patterns it will be paramount to develop methods that allow smooth transitions among the associated levels of abstraction. Formal groundings for view transformations, similar to the methods described in [6] will be required. This will need to be accompanied by other powerful visual tools, of course, in order to actually give users access to these view transformations. BioMine [8] or CET [10] have been used successfully but even more flexible methods will be needed to integrate various views within the same structure. An interesting additional challenge will be the integration of user feedback not only in terms of guiding the search but also with respect to actually learning from the users' feedback to avoid proposing uninteresting patterns over and over again. Unfortunately, as discussed above, "(un)interesting" is a moving target and heavily depends on the current scope of analysis. Active Learning approaches offer interesting mechanisms to quickly update internal models of interest to make those systems respond in a useful way. An interesting side effect could be that such learning systems observe the users, learn patterns of bisociation that were of interest in the past and actually transfer those patterns among different analyses, thus forming meta level bisociations over time.

6   Conclusions

Bisociative Knowledge Discovery promises great impact especially in those areas of scientific research where data gathering still outpaces model understanding. Once the mechanisms are well understood the task of data analysis tends to change and the focus on (statistically) significant and validated patterns is much stronger. However, in the early phase of research, the ability to collect data outperforms by far the experts' ability to make sense out of those gigantic data repositories and use them to form new hypotheses. This trend can be seen in the life sciences where data analysis barely scratches the surface of the wealth of generated data. Current methods not only fall short of offering true, explorative access to patterns within domains, but are also considerably lacking when it

comes to offering this kind of access across domains. The framework sketched here (and more substantially founded in [7]) can help to address this shortcoming. Much work still needs to be done, however, as many more types of bisociations can be formalised and many of the existing methods in the machine learning and data analysis/mining community are waiting to be applied to these problems.

One very interesting development here can be seen in the network-based bisociation discovery methods which are beginning to bridge the gap nicely between solidly understood graph theoretical algorithms and overly heuristic, poorly controllable methods. Putting those together can lead to the discovery of better understood bisociative (and other) patterns in large networks.

The data mining community has been looking for an exciting "Grand Challenge" for a number of years now. Bisociative Knowledge Discovery could offer just that: inventing methods and building systems that support the discovery of truly new knowledge across different domains will have an immense impact on how research in many fields can be computer supported in the future.

References

1. Berthold, M.R.: Bisociative Knowledge Discovery. In: Gama, J., Bradley, E., Hollmén, J. (eds.) IDA 2011. LNCS, vol. 7014, pp. 1–7. Springer, Heidelberg (2011)
2. Berthold, M.R. (ed.): Bisociative Knowledge Discovery. LNCS (LNAI), vol. 7250. Springer, Heidelberg (2012)

3. Berthold, M.R., Dill, F., Kötter, T., Thiel, K.: Supporting Creativity: Towards Associative Discovery of New Insights. In: Washio, T., Suzuki, E., Ting, K.M., Inokuchi, A. (eds.) PAKDD 2008. LNCS (LNAI), vol. 5012, pp. 14–25. Springer, Heidelberg (2008)

4. Boden, M.A.: Précis of the creative mind: Myths and mechanisms. Behavioural and Brain Sciences 17, 519–570 (1994)

5. Borgelt, C.: Network Creation: Overview. In: Berthold, M.R. (ed.) Bisociative Knowledge Discovery. LNCS (LNAI), vol. 7250, pp. 51–53. Springer, Heidelberg (2012)

6. Dries, A., Nijssen, S., De Raedt, L.: BiQL: A Query Language for Analyzing Information Networks. In: Berthold, M.R. (ed.) Bisociative Knowledge Discovery. LNCS (LNAI), vol. 7250, pp. 147–165. Springer, Heidelberg (2012)

7. Dubitzky, W., Kötter, T., Schmidt, O., Berthold, M.R.: Towards Creative Information Exploration Based on Koestler's Concept of Bisociation. In: Berthold, M.R. (ed.) Bisociative Knowledge Discovery. LNCS (LNAI), vol. 7250, pp. 11–32. Springer, Heidelberg (2012)

8. Eronen, L., Toivonen, H.: Biomine: Predicting links between biological entities using network models of heterogeneous database. BMC Bioinformatics (accepted for publication, 2012)

9. Gossen, T., Nitsche, M., Haun, S., Nürnberger, A.: Data Exploration for Knowledge Discovery: A Brief Overview of Tools and Evaluation Methods. In: Berthold, M.R. (ed.) Bisociative Knowledge Discovery. LNCS (LNAI), vol. 7250, pp. 287–300. Springer, Heidelberg (2012)

10. Haun, S., Gossen, T., Nürnberger, A., Kötter, T., Thiel, K., Berthold, M.R.: On the Integration of Graph Exploration and Data Analysis: The Creative Exploration Toolkit. In: Berthold, M.R. (ed.) Bisociative Knowledge Discovery. LNCS (LNAI), vol. 7250, pp. 301–312. Springer, Heidelberg (2012)

11. Henderson, K., Gallagher, B., Li, L., Akoglu, L., Eliassi-Rad, T., Tong, H., Faloutsos, C.: It's who you know: graph mining using recursive structural features. In: KDD, pp. 663–671 (2011)

12. Hynönen, T., Mahler, S., Toivonen, H.: Discovery of Novel Term Associations in a Document Collection. In: Berthold, M.R. (ed.) Bisociative Knowledge Discovery. LNCS (LNAI), vol. 7250, pp. 91–103. Springer, Heidelberg (2012)

13. Koestler, A.: The Act of Creation. Macmillan (1964)

14. Kötter, T., Berthold, M.R.: From Information Networks to Bisociative Information Networks. In: Berthold, M.R. (ed.) Bisociative Knowledge Discovery. LNCS (LNAI), vol. 7250, pp. 33–50. Springer, Heidelberg (2012)

15. Kötter, T., Berthold, M.R. (Missing) Concept Discovery in Heterogeneous Information Networks. In: Berthold, M.R. (ed.) Bisociative Knowledge Discovery. LNCS (LNAI), vol. 7250, pp. 230–245. Springer, Heidelberg (2012)

16. Mozetič, I., Lavrač, N.: Applications and Evaluation: Overview. In: Berthold, M.R. (ed.) Bisociative Knowledge Discovery. LNCS (LNAI), vol. 7250, pp. 359–363. Springer, Heidelberg (2012)

17. Nagel, U., Thiel, K., Kötter, T., Piatek, D., Berthold, M.R.: Towards Discovery of Subgraph Bisociations. In: Berthold, M.R. (ed.) Bisociative Knowledge Discovery. LNCS (LNAI), vol. 7250, pp. 263–284. Springer, Heidelberg (2012)

18. Nürnberger, A.: Exploration: Overview. In: Berthold, M.R. (ed.) Bisociative Knowledge Discovery. LNCS (LNAI), vol. 7250, pp. 285–286. Springer, Heidelberg (2012)

19. Rüping, S., Scheffer, T. (eds.): Learning with Multiple Views. Processings of the ICML 2005 Workshop (2005)

20. Segond, M., Borgelt, C.: Cover Similarity Based Item Set Mining. In: Berthold, M.R. (ed.) Bisociative Knowledge Discovery. LNCS (LNAI), vol. 7250, pp. 104–121. Springer, Heidelberg (2012)
21. Sluban, B., Juršič, M., Cestnik, B., Lavrač, N.: Exploring the Power of Outliers for Cross-domain Literature Mining. In: Berthold, M.R. (ed.) Bisociative Knowledge Discovery. LNCS (LNAI), vol. 7250, pp. 325–337. Springer, Heidelberg (2012)
22. Thiel, K., Berthold, M.R.: Node Similarities from Spreading Activation. In: Berthold, M.R. (ed.) Bisociative Knowledge Discovery. LNCS (LNAI), vol. 7250, pp. 246–262. Springer, Heidelberg (2012)
23. Toivonen, H.: Network Analysis: Overview. In: Berthold, M.R. (ed.) Bisociative Knowledge Discovery. LNCS (LNAI), vol. 7250, pp. 144–146. Springer, Heidelberg (2012)
24. Wiswedel, B., Hoeppner, F., Berthold, M.R.: Learning in parallel universes. Data Mining and Knowledge Discovery 21, 130–150 (2010)