Improving Transfer Learning by Introspective Reasoner

Zhongzhi Shi[1], Bo Zhang[1,2], Fuzhen Zhuang[1]

[1] Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences,100190,Beijing,China
[2] Graduate School of the Chinese Academy of Sciences, 100039, Beijing, China
{shizz, zhangb, zhuangfz}@ics.ict.ac.cn

Abstract. Traditional learning techniques have the assumption that training and test data are drawn from the same data distribution, and thus they are not suitable for dealing with the situation where new unlabeled data are obtained from fast evolving, related but different information sources. This leads to the cross-domain learning problem which targets on adapting the knowledge learned from one or more source domains to target domains. Transfer learning has made a great progress, and a lot of approaches and algorithms are presented. But negative transfer learning will cause trouble in the problem solving, which is difficult to avoid. In this paper we have proposed an introspective reasoner to overcome the negative transfer learning.

Introspective learning exploits explicit representations of its own organization and desired behavior to determine when, what, and how to learn in order to improve its own reasoning. According to the transfer learning process we will present the architecture of introspective reasoner for transductive transfer learning.

Keywords: Introspective reasoner; Transfer learning; Negative transfer

1 Introduction

Introspection method is early a psychological research approach. It investigates the psychological phenomena and process according to the report of the tested person or the experience described by himself. The introspection learning is to introduce introspection concept into machine learning. That is to say, by checking and caring about knowledge processing and reasoning method of intelligence system itself and finding out problems from failure or poor efficiency, the introspection learning forms its own learning goal and then improves the method to solving problems [1].

Cox's paper mentioned that research on introspective reasoning has a long history in artificial intelligence, psychology, and cognitive science [2]. Leake et al. pointed that in introspective learning approaches, a system exploits explicit representations of its own organization and desired behavior to determine when, what, and how to learn in order to improve its own reasoning [3]

Introspective learning has become more and more important in recent years as AI

systems have begun to address real-world problem domains, characterized by a high degree of complexity and uncertainty. Early of 1980s, introspective reasoning is implemented as planning within the meta-knowledge layer. SOAR [4] employs a form of introspective reasoning by learning meta-rules which describe how to apply rules about domain tasks and acquire knowledge. Birnbaum et al. proposed the use of self-models within case-based reasoning [5]. Cox and Ram proposed a set of general approaches to introspective reasoning and learning, automatically selecting the appropriate learning algorithms when reasoning failures arise [6]. They defined a taxonomy of causes of reasoning failures and proposed a taxonomy of learning goals, used for analyzing the traces of reasoning failures and responding to them. The main step in the introspective learning process is to looking for the reason of failure according to the features of failure. The introspective learning system should evaluate reasoning besides finding errors. The blame assignment is like troubleshooting with a mapping function from failure symptom to failure cause. A series of processes based on case-based reasoning, such as search, adjustment, evaluation and reservation can attain blame assignment and explanation failure, and improve the efficiency of evaluation and explanation, so case-based reasoning is an effective approach.

Introspective reasoning can be a useful tool for autonomously improving the performance of a CBR system by reasoning about system problem solving failures. Fox and Leake described a case-based system called ROBBIE which uses introspective reasoning to model, explain, and recover from reasoning failures [7]. Fox and Leake also took a model-based approach to recognize and repair reasoning failures. Their particular form of introspective reasoning focuses on retrieval failures and case index refinement. Work by Oehlmann, Edwards and Sleeman addressed on the related topic of re-indexing cases, through introspective questioning, to facilitate multiple viewpoints during reasoning [8]. Leake, Kinley, and Wilson described how introspective reasoning can also be used to learn adaptation knowledge in the form of adaptation cases [9]. Zhang and Yang adopted introspective learning to maintain feature weight in case-based reasoning [10]. Craw applied introspective learning to build case-based reasoning knowledge containers [11]. Introspective reasoning to repair problems may also be seen as related to the use of confidence measures for assessing the quality of the solutions proposed by a CBR system [12]. Arcos, Mulayim and Leake proposed an introspective model for autonomously improving the performance of a CBR system by reasoning about system problem solving failures [13]. The introspective reasoner monitors the reasoning process, determines the causes of the failures, and performs actions that will affect future reasoning processes. Soh et al. developed CBRMETAL system which integrates case-based reasoning and meta-learning to introspectively refine the reasoning of intelligent tutoring system [14]. CBRMETAL system specifically improves the casebase, the similarity heuristics and adaptation heuristics through reinforcement and adheres to a set of six principles to minimize interferences during meta-learning. Leake and Powell proposed WebAdapt approach which can introspectively correct failures in its adaptation process and improve its selection of Web sources to mine [15].

Transfer learning has already achieved significant success. One of the major challenges in developing transfer methods is to produce positive transfer between appro-

priately related tasks while avoiding negative transfer between tasks that are less related. So far there are some approaches for avoiding negative transfer.

Negative transfer happens when the source domain data and task contribute to the reduced performance of learning in the target domain. Despite the fact that how to avoid negative transfer is a very important issue, little research work has been published on this topic. Torrey and Shavlik summarized into three methods, that is, rejecting bad information, choosing a source task, modeling task similarity [16]. Croonenborghs et al. proposed an option-based transfer in reinforcement learning is an example of the approach that naturally incorporates the ability to reject bad information [17]. Kuhlmann and Stone looked at finding similar tasks when each task is specified in a formal language [18]. They constructed a graph to represent the elements and rules of a task. This allows them to find identical tasks by checking for graph isomorphism, and by creating minor variants of a target-task graph, they can also search for similar tasks. Eaton and DesJardins proposed to choose from among candidate solutions to a source task rather than from among candidate source tasks [19]. Carroll and Seppi developed several similarity measures for reinforcement learning tasks, comparing policies, value functions, and rewards[20].

Besides above mentioned methods, Bakker and Heskes adopted a Bayesian approach in which some of the model parameters are shared for all tasks and others more loosely connected through a joint prior distribution that can be learned from the data [21]. The data are clustered based on the task parameters, where tasks in the same cluster are supposed to be related to each other. Argyriou et al. divided the learning tasks into groups [22]. Tasks within each group are related by sharing a low-dimensional representation, which differs among different groups. As a result, tasks within a group can find it easier to transfer useful knowledge. Though there have been lots of transfer learning algorithms, few of them can guarantee to avoid negative transfer. In this paper we propose a method which adopts introspective reasoner to avoid negative transfer.

The rest of this paper is organized as follows. The following section describes transfer learning. Section 3 introduces the transductive transfer learning. Section 4 discusses introspective reasoner. Failure taxonomy in introspective reasoner is described in Section 5. Finally, the conclusions and further works are given.

2 Transfer Learning

Transfer learning is a very hot research topic in machine learning and data mining area. The study of transfer learning is motivated by the fact that people can intelligently apply knowledge learned previously to solve new problems faster or with better solutions. The fundamental motivation for transfer learning in the field of machine learning was discussed in a NIPS-95 workshop on Learning to Learn. Since 1995 Research on transfer learning has attracted more and more attention. In 2005, the Broad Agency Announcement (BAA) 05-29 of Defense Advanced Research Projects Agency (DARPA)'s Information Processing Technology Office (IPTO) gave a new mission of transfer learning: the ability of a system to recognize and apply knowledge

and skills learned in previous tasks to novel tasks. In this definition, transfer learning aims to extract the knowledge from one or more source tasks and apply the knowledge to a target task[23]. Fig. 1 shows you the transfer learning process.

The formal definitions of transfer learning are given follows. For simplicity, we only consider the case where there is one source domain D_S, and one target domain, D_T.

Definition 1 Source domain data *DS*

$$D_S = \{(x_{S_1}, y_{S_1}), \cdots (x_{S_n}, y_{S_n})\} \tag{1}$$

where $x_{S_i} \in X_S$ is the data instance and $y_{S_i} \in Y_S$ is the corresponding class label.

Definition 2 Target domain data *DT*

$$D_T = \{(x_{T_1}, y_{T_1}), \cdots (x_{T_m}, y_{T_m})\} \tag{2}$$

where the input $x_{T_i} \in X_T$ and $y_{T_i} \in Y_T$ is the corresponding output.

Definition 3 Transfer learning: Given a source domain *DS* and learning task *TS*, a target domain *DT* and learning task *TT* , transfer learning aims to help improve the learning of the target predictive function in *DT* using the knowledge in *DS* and *TS*, where *DS* ≠*DT*, and *TS* ≠ *TT* .

Definition 4 Negative transfer: In transfer learning if the source task is not sufficiently related or if the relationship is not well leveraged by the transfer method, the performance may not only fail to improve but also result in actually decreasing.
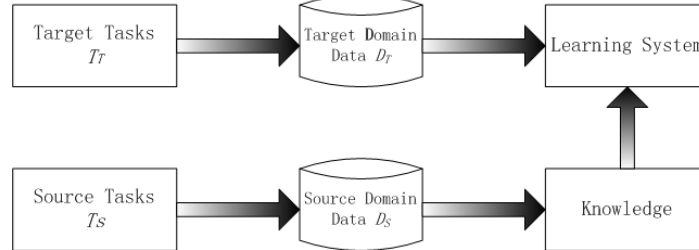

Fig. 1. Transfer learning process

The goal of our introspective reasoning system is to detect reasoning failures, refine the function of reasoning mechanisms, and improve the system performance for future problems. To achieve this goal, the introspective reasoner monitors the reasoning process, determines the possible causes of its failures, and performs actions that will affect future reasoning processes.

In this paper, we only focus on the negative transfer in the transductive transfer learning, that means, the source and target tasks are the same, while the source and target domains are different. In this situation, no labeled data in the target domain are available while a lot of labeled data in the source domain are available.

3   Transductive Transfer Learning

In the transductive transfer learning, the source and target tasks are the same, while the source and target domains are different. In this situation, no labeled data in the target domain are available while a lot of labeled data in the source domain are available. Here we give the definition of transductive transfer learning as follows:

Definition 5 Transductive transfer learning: Given a source domain *DS* and a corresponding learning task *TS*, a target domain *DT* and a corresponding learning task *TT* , transductive transfer learning aims to improve the learning of the target predictive function f_T (•) in *DT* using the knowledge in *DS* and *TS*, where $DS \neq DT$ and $TS = TT$.

   Definition 6 is clear that transductive transfer learning has different data sets and same learning tasks. In the transductive transfer learning, we want to learn an optimal model for the target domain by minimizing the expected risk [23],

$$\theta^* = \arg\min_{\theta \in \Theta} \sum_{(x,y) \in D_T} P(D_T) l(x, y, \theta)$$ (3)

where $l(x, y, \theta)$ is a loss function that depends on the parameter θ. Since no labeled data in the target domain are observed in training data, we have to learn a model from the source domain data instead. If $P(DS) = P(DT)$, then we may simply learn the model by solving the following optimization problem for use in the target domain,

$$\theta^* = \arg\min_{\theta \in \Theta} \sum_{(x,y) \in D_S} P(D_S) l(x, y, \theta)$$ (4)

If $P(DS) \neq P(DT)$, we need to modify the above optimization problem to learn a model with high generalization ability for the target domain, as follows:

$$\theta^* = \arg\min_{\theta \in \Theta} \sum_{(x,y) \in D_S} \frac{P(D_T)}{P(D_S)} P(D_S) l(x, y, \theta)$$

$$\approx \arg\min_{\theta \in \Theta} \sum_{i=1}^{n_S} \frac{P_T(x_{T_i}, y_{T_i})}{P_S(x_{S_i}, y_{S_i})} l(x_{S_i}, y_{S_i}, \theta)$$ (5)

By adding different penalty values to each instance (x_{Si}, y_{Si}) with the corresponding weight $\frac{P_T(x_{T_i}, y_{T_i})}{P_S(x_{S_i}, y_{S_i})}$, we can learn a precise model for the target domain. Since $P(YT|XT) = P(YS|XS)$, the differences between *P(DS)* and *P(DT)* are caused by *P(XS)* and *P(XT)* and

$$\frac{P_T(x_{T_i}, y_{T_i})}{P_S(x_{S_i}, y_{S_i})} = \frac{P(x_{T_i})}{P(x_{S_i})}$$

If we can estimate $\frac{P(x_{S_i})}{P(x_{T_i})}$ for each instance, we can solve the transductive transfer learning problems.

In transductive transfer learning the distributions on source domain and target domain do not match, we are facing sample selection bias or covariate shift. Specifically, given a domain of patterns X and labels Y, we obtain source training samples $Z_S=\{(x_{S_1}, y_{S_1}), \ldots, (x_{S_n}, y_{S_n})\} \subseteq X \times Y$ from a Borel probability distribution $PS(x, y)$, and target samples $Z_T=\{(x_{T_1}, y_{T_1}), \ldots, (x_{T_m}, y_{T_m})\}$ drawn from another such distribution $P_T(x, y)$. Huang et al. proposed a kernel-mean matching (KMM) algorithm to learn $\frac{P(x_{S_i})}{P(x_{T_i})}$ directly by matching the means between the source domain data and the target domain data in a reproducing-kernel Hilbert space (RKHS) [24]. KMM can be rewritten as the following optimization problem.

$$\min_{\beta} \frac{1}{2} \beta^T K \beta - k^T \beta$$

$$\text{s.t.} \quad \beta_i \in [0, B] \text{ and } | \sum_{i=1}^{n_S} \beta_i - n_S | \leq n_S \varepsilon$$

(6)

where

$$K = \begin{bmatrix} K_{S,S} & K_{S,T} \\ K_{T,S} & K_{T,T} \end{bmatrix}$$

and $K_{i,j}=k(x_i, x_j)$, $K_{S,S}$ and $K_{T,T}$ are kernel matrices for the source domain data and the target domain data, respectively. Note that (6) is a quadratic program which can be solved efficiently using interior point methods or any other successive optimization procedure. Huang et al. proved that

$$P(x_{T_i}) = \beta_i P(x_{S_i})$$

(7)

The advantage of using KMM is that it can avoid performing density estimation of either $P(x_{Si})$ or $P(x_{Ti})$, which is difficult when the size of the data set is small.

A common assumption in supervised learning is that training and test samples follow the same distribution. However, this basic assumption is often violated in practice and then standard machine learning methods do not work as desired. The situation where the input distribution $P(x)$ is different in the training and test phases but the conditional distribution of output values $P(y/x)$ remains unchanged is called covariate shift. Sugiyama et al. proposed an algorithm known as Kullback-Leibler Importance Estimation Procedure(KLIEP) to estimate $\frac{P(x_{S_i})}{P(x_{T_i})}$ directly[25], based on the minimization of the Kullback-Leibler divergence. The optimization criterion is as follows:

$$\max_{\{\theta_l\}_{l=1}^{b}} \left[\sum_{j=1}^{m} \log \sum_{l=1}^{b} \theta_l \phi_l(\mathbf{x}_j^m) \right]$$

$$\text{subject to } \sum_{i=1}^{n} \sum_{l=1}^{b} \theta_l \phi_l(x_i^n) \tag{8}$$

It can be integrated with cross-validation to perform model selection automatically in two steps: a) estimating the weights of the source domain data and b) training models on the reweighted data.

The Naïve Bayes classifier is effective for text categorization [26]. Regarding the text categorization problem, a document $d \in D$ corresponds to a data instance, where D denotes the training document set. The document d can be represented as a bag of words. Each word $w \in d$ comes from a set W of all feature words. Each document d is associated with a class label $c \in C$, where C denotes the class label set. The Naïve Bayes classifiers estimate the conditional probability as follows:

$$P(c \mid d) = P(c) \prod_{w \in d} P(w \mid c) \tag{9}$$

Dai et al. proposed a transfer-learning algorithm for text classification based on an EM-based Naïve Bayesian classifiers NBTC [27]. Dai et al. also proposed co-clustering based classification for out-of-domain documents algorithm CoCC which takes co-clustering as a bridge to propagate the knowledge from the in-domain to out-of-domain [28].

Ling et al. proposed a spectral classification based method CDSC, in which the labeled data from the source domains are available for training and the unlabeled data from target domains are to be classified [29]. Based on the normalized cut cost function, supervisory knowledge is transferred through a constraint matrix, and the regularized objective function finds the consistency between the source domain supervision and the target domain intrinsic structure. Zhuang et al. proposed CCR_3 approach which is a consensus regularization framework to exploit the distribution differences and learn the knowledge among training data from multiple source domains to boost the learning performance in a target domain [30].

4   Introspective Reasoner

The goal of our introspective reasoning system is to detect negative transfer , refine the function of reasoning mechanisms, and improve the system performance for future problems. In order to reach this goal, the introspective reasoner should solve following problems [31]:

1. There are standards that determine when the reasoning process should be checked, i.e. monitoring the reasoning process;
2. Determine whether failure reasoning takes place according to the standards;

3. Confirm the final reason that leads to the failure;
4. Change the reasoning process in order to avoid the similar failure in the future.

Introspective reasoner for transfer learning will monitor the transfer learning process, determine the possible causes of its failures, and perform actions that will affect future transfer processes. The architecture of introspective reasoner is shown in Fig. 2.
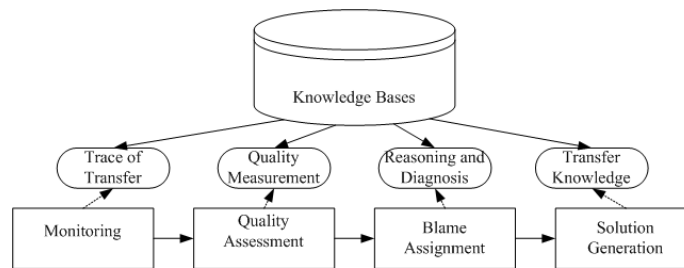

Fig. 2. Architecture of Introspective Reasoner

Introspective reasoner consists of monitoring, quality assessment, blame assignment, solution generation and ontology-based knowledge bases. The monitoring task tracks the transductive transfer learning process. For each learning task solved by the transductive transfer learning system, the monitor generates a trace containing: a) the source data retrieved; b) the ranking criteria applied to the data, together with the values that each criterion produced and the final ranking; and c) the transfer operators which were applied, with the sources to which they were applied and the target changes produced.

When the user's final solution is provided to the system, quality assessment is triggered to determine the real quality of the system-generated solution, by analyzing the differences between the system's proposed solution and the final solution. Quality assessment provides a result in qualitative terms: positive quality or negative quality.

Blame assignment starts by negative quality assessment. It takes as input the differences between the source data and target data, and tries to relate the solution differences to the source data. The system searches the ontology-based knowledge bases and selects those that apply to the observed solution differences.

Solution generation identifies the learning goals related to the negative transfer failure selected in the blame assignment stage. Each failure may be associated with more than one learning goal. For each learning goal, a set of plausible source data changes in the active policies are generated by using a predefined ontology-based knowledge base.

Failure taxonomy is an important problem in introspective learning system. It can provide a clue for the failure explanation and the revise learning target. In the introspective reasoner system, the failures are put into knowledge bases which use ontology as knowledge representation. In this way, the knowledge is organized as conceptualization, formal, semantics explicit and shared [32].

5   Failure Taxonomy In Introspective Reasoner

Failure taxonomy is an important problem in introspective learning system. Hierar-chical failure taxonomy is a method that can solve the problem classification is too fine or too coarse in failure taxonomy. Failure can be represented into large class according to the different reasoning step. This can improve the introspective revision of the system and accelerate the process of failure comparing. If failure taxonomy is fine, failures can be described more clearly, in order to provide valuable clue. Giving a rational treatment for the relationship between the failure taxonomy and failure interpretation will increase the capacity of the system introspection. System not only need to reason out the reasons of failure and give introspective learning target accord to the symptoms of failure, but also need to have all capacities to deal with different problems. Similarly, failure explanation also can be classed into different layers, such as abstract level, level of detail or multidimensional. The layers of failure taxonomy contribute to form the reasonable relationship between the features of failure and the failure explanation.

In introspective reasoner knowledge bases contain failure taxonomy and special knowledge coming from previous experiences or expertise. Solving the failure tax-onomy in introspective learning through ontology-based knowledge bases can make the failure taxonomy in introspective learning more clear and retrieval process more efficient.
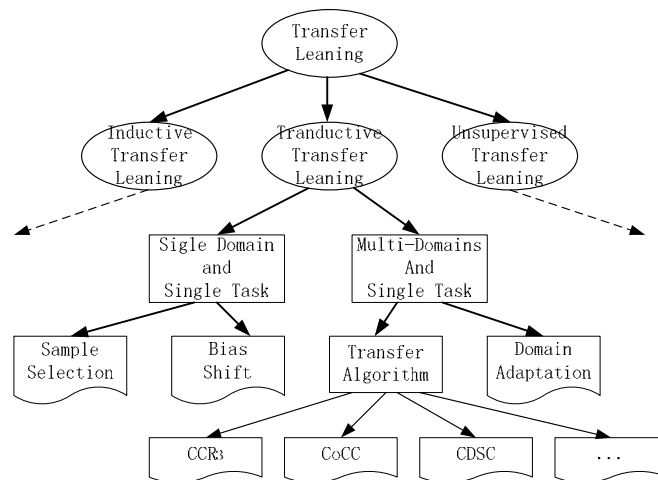

Fig. 3. Failure Taxonomy In Introspective Reasoner

Based on different situations between the source and target domains and tasks, transfer learning can be categorized three types, such as inductive transfer learning, transductive transfer learning and unsupervised transfer learning. According to the basic situation, the failure taxonomy in introspective reasoner is constructed as Fig. 3. In this paper we only focus on the transductive transfer learning which can be divided two categories, that is, single source domain single task and multi-domain single task.

At the present time, researchers make a lot approaches and algorithms in multi-domain multi task. And also domain adaptation is the key problem in the transductive transfer learning. When transfer learning failure happens we can do blame assignment.

Here we make an example from domain adaptation. Domain adaptation arises in a variety of modern applications where limited or no labeled data is available for a target application. Obtaining a good adaptation model requires the careful modeling of the relationship between $P(D_S)$ and $P(D_T)$. If these two distributions are independent, then the target domain data DT is useless for building a model. On the other hand, if $P(D_S)$ and $P(D_T)$ are identical, then there is no adaptation necessary and we can simply use a standard learning algorithm. In practical problems $P(D_S)$ and $P(D_T)$ are neither identical nor independent.

In text classification, Daume III et al. proposed the Maximum Entropy Genre Adaptation Model, which is called MEGA Model [33]. The failure taxonomy can be illustrated in Table 1. When the transductive transfer learning is failed, the diagnosis is provided as input for the blame assessment with failure categories used to suggest points to avoid negative transfer or low performance transfer.

Table 1. Sample possible sources of failure

Failed knowledge goal	Failure point	Explanation
Only target domain	No knowledge	Source domain choice
Only source domain	No transfer need	Target domain choice
Extra source domain data need	Parameter is on held out	Source domain data not enough
Adapted source domain	Negative transfer	Source domain selection

6   Conclusions

Traditional learning techniques have the assumption that training and test data are drawn from the same data distribution, and thus they are not suitable for dealing with the situation where new unlabeled data are obtained from fast evolving, related but different information sources. This leads to the cross-domain learning problem which targets on adapting the knowledge learned from one or more source domains to target domains. Transfer learning has made a great progress and a lot of approaches and algorithms are presented. But negative transfer learning will cause trouble in the problem solving and it is difficult to avoid. In this paper we have proposed an introspective reasoner to overcome the negative transfer learning.

Introspective learning exploits explicit representations of its own organization and desired behavior to determine when, what, and how to learn in order to improve its own reasoning. According to the transfer learning process, we presented the architecture of introspective reasoner with four components, which is monitoring, quality assessment, blame assignment and solution generation. All of these components are supported by knowledge bases. The knowledge bases are organized in terms of ontol-

ogy. This paper only focuses on transductive transfer learning. In the future we are working on introspective reasoner for multi-source and multi-tasks transfer learning.

Acknowledgements

References

1. Zhongzhi Shi: Intelligence Science. World Scientific, Singapore,   2011
2. Cox, M. 2005. Metacognition in computation: A selected research review. Artificial Intelligence 169(2):104–141.
3. Leake, D. B., Wilson, M.: Extending introspective learning from self- models. In M. T. Cox & A. Raja (Eds.), Metareasoning: Thinking about thinking, 2008
4. Laird, J. E., Newell, A., and Rosenbloom, P. S. Soar: A Architecture for General Intelligence. Artificial Intelligence. 1987
5. Birnbaum, L.,Collins, G., Brand, M., Freed, M., Krulwich, B., and Pryor, L.: A model-based approach to the construction of adaptive case-based planning systems. In Bareiss, R., ed., Proceedings of the DARPA Case-Based Reasoning Workshop, 215-224. San Mateo: Morgan Kaufmann,1991
6. Cox, M.T., Ram, A.: Introspective multistrategy learning on the construction of learning strategies. Artificial Intelligence, 1999, 112: 1-55
7. Fox, S. and Leake, D. B.: Using Introspective Reasoning to Refine Indexing. Proceedings of the 14th International Joint Conference on Artificial Intelligence, 1995, pp. 391-397
8. Oehlmann, R., Edwards, P., & Sleeman, D. Changing the Viewpoint: Re-Indexing by Introspective Question. Proceedings of the 16th Annual Conference of the Cognitive Science Society, 1995, pp. 381-386
9. Leake, D. B., Kinley, A., and Wilson, D. Learning to Improve Case Adaptation by Introspective Reasoning and CBR. Case-Based Reasoning Research and Development (Ed.s M. Veloso & A. Aamodt), Proceedings of the 1st International Conference on Case-Based Reasoning, 1995, pp. 229-240
10. Zhong Zhang and Qiang Yang. Feature Weight Maintenance in Case Bases Using Introspective Learning. Journal of Intelligent Information Systems, 16(2): 95-116 , 2001
11. Susan Craw: Introspective Learning to Build Case-Based Reasoning (CBR) Knowledge Containers. MLDM 2003: 1-6
12. Cheetham, W., and Price, J.: Measures of solution accuracy in case-based reasoning systems. In Funk, P., and Conzalez-Calero, P. A., eds., Proceedings of the 7th European Conference on Case-Based Reasoning, number 3155 in Lecture Notes in Artificial Intelligence. Springer-Verlag. 106-118, 2004
13. Arcos, J. L., Mulayim, O., Leake, D.: Using Introspective Reasoning to Improve CBR System Performance. *AAAI Metareasoning Workshop*: AAAI Press, pp. 21-28, 2008
14. Leen-Kiat Soh, Todd Blank: Integrated Introspective Case-Based Reasoning for Intelligent Tutoring Systems. Association for the Advancement of Artificial Intelligence. AAAI-2007

15. Leake, D. B., Powell, J. H.: Enhancing Case Adaptation with Introspective Reasoning and Web Mining. IJCAI 2011: 2680-2685, 2011
16. Lisa Torrey and Jude Shavlik: Transfer Learning. Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques, pp. 242-264, 2010
17. Croonenborghs, T., Driessens, K. and M. Bruynooghe. Learning relational skills for inductive transfer in relational reinforcement learning. In International Conference on Inductive Logic Programming, 2007.
18. Kuhlmann, G. and Stone, P.: Graph-based domain mapping for transfer learning in general games. In European Conference on Machine Learning, 2007
19. Eaton, E. and DesJardins M.: Knowledge transfer with a multiresolution ensemble learning of classifiers. In ICML Workshop on Structural Knowledge Transfer for Machine, 2006
20. Carroll, C. and Seppi, K.: Task similarity measures for transfer in reinforcement works learning task libraries. In IEEE International Joint Conference on Neural Net, 2005
21. Bakker, B. and Heskes, T.: Task Clustering and Gating for Bayesian Multitask Learning. J. Machine Learning Research, vol. 4, pp. 83-99, 2003
22. Argyriou, A., Maurer, A. and Pontil, M.: An Algorithm for Transfer Learning in a Heterogeneous Environment. Proc European Conf. Machine Learning and Knowledge Discovery in Databases (ECML/PKDD '08), pp. 71-85, Sept. 2008
23. Sinno Jialin Pan, Qiang Yang: A Survey on Transfer Learning. IEEE Trans. Knowl. Data Eng. 22(10): 1345-1359, 2010
24. Huang, J., Smola, A., Gretton, A., Borgwardt, K.M. and Scholkopf, B.: Correcting Sample Selection Bias by Unlabeled Data. Proc. 19th Ann. Conf. Neural Information Processing Systems, 2006
25. M. Sugiyama, S. Nakajima, H. Kashima, P.V. Buenau, and M. Kawanabe, Direct Importance Estimation with Model Selection and its Application to Covariate Shift Adaptation. Proc. 20th Ann. Conf. Neural Information Processing Systems, Dec. 2008
26. Lewis, D. D.: Representation and learning in information retrieval. Ph.D. Dissertation, Amherst, MA, USA, 1992
27. W. Dai, G. Xue, Q. Yang, and Y. Yu: Transferring Naive Bayes Classifiers for Text Classification. Proc. 22nd Assoc. for the Advancement of Artificial Intelligence (AAAI) Conf. Artificial Intelligence, pp. 540-545, July 2007
28. Dai W Y, Xue G R, Yang Q, et al. Co-clustering based Classification for Out-of-domain Documents [C]. // *Proceedings of 13th ACM International Conference on Knowledge Discovery and Data Mining,* New York: ACM Press, 2007: 210—219
29. Ling X, Dai W Y, Xue G R, et al. Spectral Domain-Transfer Learning [C]. // *Proceedings of 14th ACM International Conference on Knowledge Discovery and Data Mining*, New York: ACM Press, 2008: 488—496
30. Zhuang Fuzhen, Luo Ping, Xiong Hui, Xiong Yuhong, He Qing, Shi Zhongzhi. Cross-domain Learning from Multiple Sources: A Consensus Regularization Perspective [J]. *IEEE Transactions on Knowledge and Data Engineering* (TKDE), 2010, 22(12): 1664-1678, 2010
31. Zhongzhi Shi and Sulan Zhang. Case-Based Introspective Learning. Proceedings of the Fourth IEEE International Conference on Cognitive Informatics. IEEE-2005
32. Qi Dong, Zhongzhi Shi: A Research on Introspective Learning Based on CBR. International Journal of Advanced Intelligence. Volume 3, Number 1, pp.147-157, March, 2011
33. Daume III, H. and Marcu, D.: Domain Adaptation for Statistical Classifiers. J. Artificial Intelligence Research, vol. 26, pp. 101-126, 2006.