Towards an affective computational model for machine consciousness

Rohitash Chandra

Centre for Translational Data Science
The University of Sydney, NSW 2006, Australia.
rohitash.chandra@sydney.edu.au

Abstract. In the past, computational models for machine consciousness have been proposed with varying degrees of challenges for implementation. Affective computing focuses on the development of systems that can simulate, recognize, and process human affects which refer to the experience of feeling or emotion. The affective attributes are important factors for the future of machine consciousness with the rise of technologies that can assist humans and also build trustworthy relationships between humans and artificial systems. In this paper, an affective computational model for machine consciousnesses with a system of management of the major features. Real-world scenarios are presented to further illustrate the functionality of the model and provide a road-map for computational implementation.

Keywords: Affective computing, machine consciousness, learning algorithms, neuro-psychology.

1 Introduction

Throughout modern digital history, there have been a number of developments in areas of artificial intelligence that mimic aspects or attributes of cognition and consciousness. These developments have been made with the hope to replicate and automate some of the tasks that are undertaken by humans given the industrial demand and constraints of humans on carrying out demanding tasks in limited time. In an attempt to empirically study consciousness, Tononi proposed the information integrated theory of consciousness to quantify the amount of integrated information an entity possesses which determines its level of consciousness [32]. The theory depends exclusively on the ability of a system to integrate information, regardless of having a strong sense of self, language, emotion, body, or an environment. David Chalmers highlighted the explanatory gap in defining consciousness and indicated that the hard problem of consciousness emerge from attempts that try to explain it in purely physical terms [5]. Integrated information theory is based on phenomenological axioms which *begins with consciousness* and indicates that complex systems with some feedback states could have varying levels of consciousness [23]. Howsoever, this does not support the concept of conscious experience. Chalmers argued that the science of consciousness must integrate third-person data about behavior and brain processes with first-person data about conscious experience [7]. Some examples include comparing conscious and unconscious processes, investigating the contents of consciousness

[24], finding neural correlates of consciousness [6], and connecting consciousness with physical processes.

The field of affective computing focuses on the development of systems that can simulate, recognize, and process human affects which essentially is the experience of feeling or emotion [26,27]. Affective computing could provide better communication between humans and artificial systems that can lead to elements of trust and connectivity with artificial systems [31]. The motivation to have affective models in artificial consciousness would be towards the future of mobile technologies and robotic systems that guide in everyday human activities. For instance, a robotic system which is part of the household kitchen could further feature communication that builds and connectivity from features of affective computing [33]. In the near future, there will also be a growing demand for sex robots, therapeutic and nursing robots which would need affective computing features [1,29]. Moreover, the emergence of smart toys and robotic pets could be helpful in raising children and also assist the elderly [29]. Although mobile application-based support and learning systems have been successfully deployed, they are often criticized for having less physical interactions [22]. In such areas, affects in robots could lead to further help such as stress management and counselling. Hence, the affective attributes are important factors for the future of machine consciousness with the rise of technologies that can assist humans and also build trustworthy relationships with them.

In this paper, an affective computational model for machine consciousnesses is presented that features an algorithm for management of the major aspects of consciousness that range from information processing to critical thinking. Real-world problem scenarios is presented to further illustrate the functionality of the model and a road-map for software-based implementation has been also discussed.

The rest of the paper is organised as follows. Section 2 provides related work and Section 3 presents the affective computational model for machine consciousness. Section 4 provides a discussion while Section 5 concludes the paper with further research directions.

2 Related Work

The field of natural language processing aims to make computer systems understand and manipulate natural languages to perform the desired tasks [9]. It has been one of the major attributes of cognition and consciousness [15]. One of the major breakthroughs that used natural language processing for cognitive computing has been the design of Watson, which is a system capable of answering questions posed in natural language developed by David Ferrucci [10]. Watson won the game of Jeopardy against human players [19]. It had access to 200 million pages of structured and unstructured content including the full text of Wikipedia. Moreover, IBM Watson was not connected to the Internet during the game. There are a number of applications of Watson technology that includes various forms of search that have semantic properties [11]. Furthermore, it can help in developing breakthrough research in medical and life sciences with a further focus on Big Data challenges. Hence, it was shown that Watson can accelerate the identification of novel drug candidates and novel drug targets by harnessing the potential of big data [8].

With such a breakthrough for development of Watson for cognitive computing, there remains deep philosophical questions from perspective of natural and artificial consciousness [18]. Koch evaluated Watson's level of consciousnesses from perspective of integrated information theory of consciousnesses [17] that views the level of consciousness based on complexity and how integrated the forms of information are in the system. Watsons capabilities motivated to further study the philosophy, theory, and future of artificial intelligence based upon Leibnizs computational formal logic that inspired a 'scorecard' approach to assessing cognitive systems [3]. Metacognition refers to a higher order thinking skills that includes knowledge about when and how to use particular strategies for learning or for problem solving [12]. In relation to metacognition, Watson relied on a skill very similar to human self-knowledge as it not only came up with answers but also generated a confidence rating for them. Therefore, Watson possessed elements of metacognition similar to the human counterparts in the game of Jeopardy [13]. More recently, *AlphaGo* was developed by *Google* to play the board game *Go* which became the first program to beat a professional human player without handicaps on a full-sized 19×19 board [14]. Although AlphaGo has been very successful, one can argue that it demonstrated a very constrained aspect of human intelligence that may not necessarily display consciousness.

3 An Affective Computational Model

3.1 Problem Scenarios

The details of the affective model is presented with problem scenarios that are intended to demonstrate its effectiveness. Depending on the experience, there is an expression which would be involuntarily stored as either long or short-term memory depending on the nature of the experience. Moreover, there is also conceptual understanding of implications to the observer and how it changes their long and short-time goals. The output in terms of action or expression could also be either voluntary or involuntary. In some situations, one reacts without controlling their emotions while in others, one does not react. A conscious decision is made depending on the type of personality, depth of knowledge (machine learning models) from past experience (audio, visual and other data). Figure 1 shows a general view of state-based information processing based on experience which acts as input or action while the response acts as the reaction given by the behavior or expression. Figure 2 shows an over overview of the affective model of consciousness that is inter-related with Figure 1. The states in Figure 2 shown in blue represent the metaphysical while those in black are the physical states.

The accounts of situations is further presented that require problem-solving skills feature different states of consciousness. The description of the scenario and how it will be tackled by the proposed affective model is presented with three distinct scenarios as follows.

Scenario 1: *Ramon is traveling on a flight from India to Japan and has a connecting flight from Shanghai, China. His flight lands in Shanghai and he is required to make it to the connecting flight gate. Ramon's boarding pass has gate*

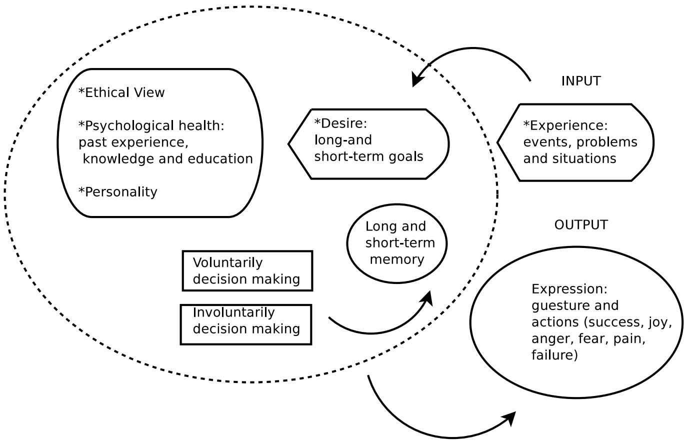

Fig. 1. Response after processing the features that contribute to a system that replicates elements of consciousness

information missing and since his flight landed about and hour late, he needs to rush to the connecting gate. Ramon is not sure if he will pass through the immigration authority. His major goal is to reach a connecting flight gate. In doing so, he is required to gather information about his gate and whether he will go through the immigration processing counter. He encounters a series of emotions which includes fear of losing the connecting flight and hence exhibits a number of actions that show his emotive psycho-physical states which include sweating, exaggerating while speaking and even shivering due to fear.

In order for Ramon to successfully make it to the connecting flight on time, he will undergo a series of states in consciousnesses which is described in detail with state references from Figure 2 as follows.

1. Exit flight and find the way to transfer desk.
   (a) Search for information regarding "transfers and arrivals" through vision recognition system (State 2 and then State 6).
   (b) Process information and make decision to move to the area of "transfers" (State 2 and State 5).
2. Since information that no baggage needs to be collected was already given, check boarding pass for baggage tag sticker.
   (a) Process visual information by checking boarding pass (State 2 and 6)
3. Confirm with the officer at transfer desk if there is a need to go through the Immigration Counter.
   (a) Find and walk to transfer desk (State 2, 6, and 5)
   (b) Communicate with the officer at transfer desk (State 2 and 6)
   (c) Fear and emotions during communication (State 2, 5, 8, and 10)
4. Information was given by the officer that there is a need to go through immigration booth, hence, prepare boarding pass and passport.
   (a) Rush to the immigration processing section (State 5 and 6).
   (b) Wait in queue and go through a number of emotions such as fear of losing flight and also sweat (State 5, 6, 8, and 10).

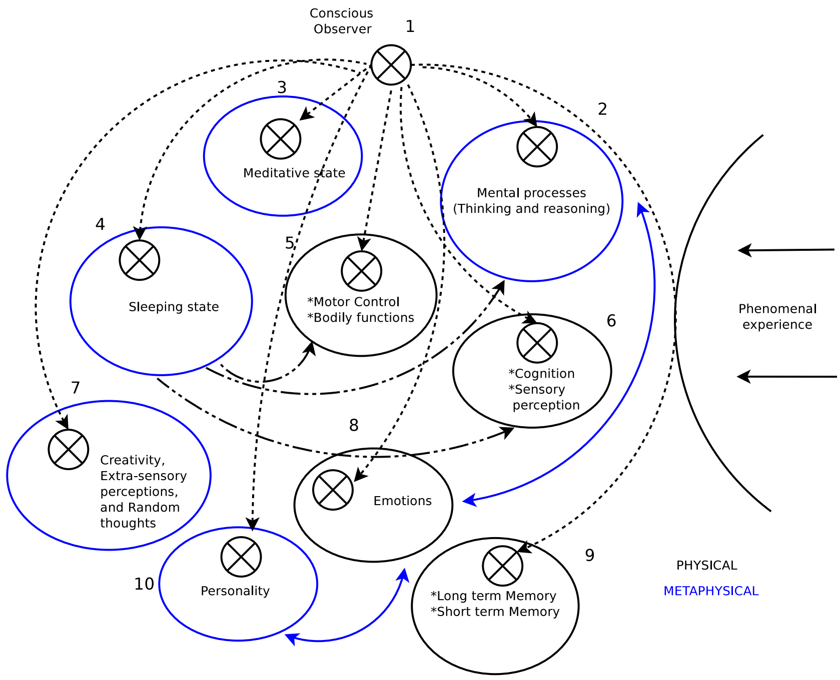

Fig. 2. The conscious observer is defined as the root of consciousness which is also referred to as "qualia". It can enter different states while also having the property to exist within two states, i.e it can self-replicate as a process, gather knowledge and update long and short-term memories. The blue states are metaphysical and black states are physical.

5. After immigration processing, find gate information and move to gate and board connecting flight.
   (a) Rush to the gate. In the process breath heavily and also sweat (State 2, 5, and 6).
   (b) Wait at the gate with some random thoughts and then board when called (State 7, 8, 6, 2 and 5).

Scenario 2: *Thomas is in a mall in Singapore for his regular Saturday movies and shopping with friends. Suddenly, he realizes that he can't locate his phone. He tries to brainstorm about the moments when he used his phone. He goes through a series of intense emotive states that includes fear.*

In order for Thomas to successfully find his phone, he will undergo a series of states in consciousnesses with reference from Figure 2 as follows.

1. Thomas first informed his friends and began checking all his pockets and carry bag.
   (a) Check all pockets (State 5 and 6).
   (b) Inform friends and also check in carry bag (State 6, 8, 10, and 5)
2. Brainstorm where was last time phone was used.
   (a) Ask friends if they can recall him using the phone (State 6, 8, and 10).

   (b) Try to remember when phone was last used (State 2 and 9).
   (c) Finally, take a moment of a deep breath and relax in order to remember (State 2, 9, and 3).
3. Recalled from memory that phone was last used in cinema and then rush there to check.
   (a) Recalled that phone was last used in cinema (State 1, 3, 7, and 9).
   (b) Inform friends with emotive expression of hope and achievement (State 6 and 8).
   (c) Rush to the cinema and talk to the attendant with emotive state of hope and fear (State 5, 6 and 8).
   (d) Attendant locates the phone and informs (State 6).
   (e) Emotive state of joy and achievement (State 8).

3.2 Artificial Qualia Manager

The affective computational model has the potential to replicate elements of human consciousness. It can exhibit characteristics with human touch with emotive states through synergy with affective computing. There is a need for management of components in the affective model which would help the property of conscious experience. Hence, there is a need for a manager for qualia. This could be seen as a root algorithm that manages the states with features that can assign the states based on the goal and the needs (instincts) and qualities (such as personality and knowledge).

The artificial qualia manager could be developed with the underlying principle in the case of a security guard officer that monitors a number of video feedbacks from security cameras and uses radio communication with other security guards. The officer follows a channel of communication strategies if any risks or security impeachment occurs. The artificial qualia manager would be monitoring and managing the states to assign jobs for reaching the goal through automated reasoning in machine consciousness as given in Algorithm 1.

In Algorithm 1, the goal and data from audio and visual inputs are used to determine and effectively manage the sequence of states of affective model of consciousness presented in Figure 2. Once the goal is reached, a series of states can be used for expression which can include a set of emotions. Note that audio and visual data needs to undergo through processing with machine learning tools which would then output some information. For instance, if the goal is regarding finding date information for a boarding pass, then the task would be to be first to translate this higher level task into a sequence of lower level tasks that would execute machine learning components. After these components are triggered, they would return information which will be used by the algorithm to make further decision of states needed to reach the goal. This is illustrated in Figure 3. There needs to a be a property of states for tasks based on their importance. For instance, priority is given to emergency situations while trying to fill a goal. While fulfilling a goal, priority would be given to aspects such as safety and security. The goal could be similar to those given in Scenario 1 and Scenario 2 where Ramon boards connecting flight and Thomas locates his phone, respectively.

Alg. 1 Artificial Qualia Manager

Data: Data from sensory perception (video, audio, and sensor data)
Result: States for consciousness
Initialization (knowledge and personality)

Φ is the list of states
Σ is the final gaol to reach
Ω is the list of actions required to reach the goal

```
while alive do
    traverse-states(Σ, Φ])

    while goal Σ not reached do
        if challenge then
            1. Nominate a state
            2. Attend to challenge (injury, pain, emotion)
            3. Store short-term and long-term memory
        end
        if goal Σ reached (success) then
            1. Output through expression (action, gesture, emotion)
            2. Store short-term and long-term memory
        end
        if goal Σ not reached (failure) then
            1. Output through expression (action, gesture, emotion)
            2. Store short-term and long-term memory
        end
    end
    1. Generate random thoughts based on problem and emotion
    2. Automated reasoning and planning for states needed for future goal(s)
    3. Address the requirements to revisit failed goals
end
```

3.3 Implementation Strategies

Multi-task learning is motivated by cognitive behavior where the underlying knowledge from one task is helpful to one or several other tasks. Hence, multi-task learning employs sharing of fundamental knowledge across tasks [4,25]. The affective computational model can feature multi-task learning for replicating sensory perception through recognition task that includes vision, sensory input for touch and smell and auditory tasks such as speech verification, speech recognition, and speaker verification. Shared knowledge representation would further be used for recognition of objects, faces or facial expression where visual and auditory signals would be used in conjunction to make a decision.

In the identification of objects, its is important learn through the experience of different senses that can be seen as a modular input to biological neural system [16]. Modular learning would help in decision making in cases where one of the signals is not available [30]. For instance, a humanoid robot is required to recognize someone in the dark when no visual signal is available, it would be able to make a decision based on the auditory signal. Ensemble learning could take advantage of several machine learning models which can also include deep learning for visual or auditory based recognition systems [28]. Ensemble learning can also be used to address multi-label learning where instances have multiple labels which is different from multi-class learning [34]. Furthermore, language models that feature neural networks [2,21] could be used in conjunction with speech recognition methods [20].

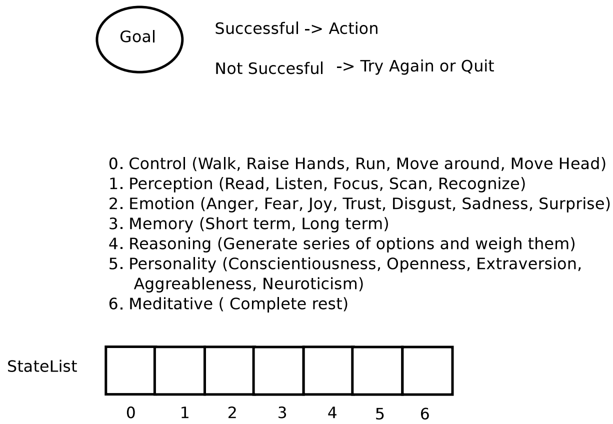

Fig. 3. States in affective computational for the artificial qualia manager

4 Discussion

It is important to highlight the potential of animal consciousness as it can motivate models for consciousness that fill the gaps in models for human consciousness. In simulation or the need to implant certain level of consciousnesses to robotic systems, it would be reasonable to begin with animal level where certain tasks can be achieved. For instance, a robotic system that can replicate cognitive abilities and level of consciousness for rats can be used for some tasks such as burrowing holes, navigation in unconstrained areas for feedback of videos or information, in disasters such as earthquakes and exploration of remote places, and evacuation sites.

Deep learning, data science and analytics can further help in contribution towards certain or very limited areas of machine consciousness. This is primary to artificially replicate areas of sensory input such as artificial speech recognition and artificial vision or perception. Howsoever, with such advancements in replication of sensory perceptions, one encounters further challenges in developing software systems that manage aspects of perception that lead to a conscious state. With the rise of technologies such as Internet of Things (IoT), sensors could be used to replicate biological attributes such as pain, emotions, feeling of strength and tiredness. However, modelling these attributes and attaining same behaviour in humans may not necessarily mean that the affective model would address hard problem of consciousness. However, at least the model would be seen to replicate conscious experience to a certain degree that could be similar to humans and other animals. Such an affective model, with future implementations could give rise to household robotic pets that would have or could develop emotional relationship with humans. The concern would in giving autonomous control or decision making through simulated aspects of emotion and human behaviour. Humans are well known to be poor decision makers when in emotional states which also resort to level of aggression and violence. Therefore, simulation of affective states need to take into account of safety and security for the future robots that assist humans.

5 Conclusions and Future Work

The paper presented the notion of using affects in computational model for machine consciousness in order to give a human-like expression or behavior for artificial systems. The challenges lie in further refining specific features such as personality and creativity. Howsoever, the proposed model can be a baseline and motivate the coming decade of simulation and implementation of machine consciousness for artificial systems such as humanoid robots. The simulation for affective model of consciousness with the features of artificial qualia manager can also be implemented with the use of robotics hardware. In their absence, simulation can also be implemented through collection of audiovisual data and definition of certain goals. The affective model is general and does not only apply to humanoid robots, but can be implemented in service application areas of software systems and technology.

Future research could concentrate in simulation of the proposed model and development of areas such as artificial personality in machine consciousness. The incorporation of technologies such as IoT, semantic web, cognitive computing and machine learning could guide in simulation of aspects of machine consciousness.

References

1. Bendel, O.: Surgical, Therapeutic, Nursing and Sex Robots in Machine and Information Ethics, pp. 17–32. Springer International Publishing, Cham (2015)
2. Bengio, Y., Ducharme, R., Vincent, P., Jauvin, C.: A neural probabilistic language model. Journal of machine learning research 3(Feb), 1137–1155 (2003)
3. Bringsjord, S., Govindarajulu, N.S.: Leibniz's Art of Infallibility, Watson, and the Philosophy, Theory, and Future of AI, pp. 183–200. Springer International Publishing, Cham (2016)
4. Caruana, R.: Multitask learning. Machine Learning 28(1), 41–75 (Jul 1997)
5. Chalmers, D.J.: Facing up to the problem of consciousness. Journal of Consciousness Studies 2(3), 200–19 (1995)
6. Chalmers, D.J.: What is a neural correlate of consciousness. Neural correlates of consciousness: Empirical and conceptual questions pp. 17–40 (2000)
7. Chalmers, D.J.: How can we construct a science of consciousness? Annals of the New York Academy of Sciences 1303(1), 25–35 (2013)
8. Chen, Y., Argentinis, J.E., Weber, G.: {IBM} watson: How cognitive computing can be applied to big data challenges in life sciences research. Clinical Therapeutics 38(4), 688 – 701 (2016)
9. Chowdhury, G.G.: Natural language processing. Annual review of information science and technology 37(1), 51–89 (2003)
10. Ferrucci, D.: Build watson: an overview of deepqa for the jeopardy! challenge. In: Proceedings of the 19th international conference on Parallel architectures and compilation techniques. pp. 1–2. ACM (2010)
11. Ferrucci, D., Levas, A., Bagchi, S., Gondek, D., Mueller, E.T.: Watson: Beyond jeopardy! Artificial Intelligence 199, 93 – 105 (2013)

12. Flavell, J.H.: Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. American psychologist 34(10), 906 (1979)
13. Fleming, S.M.: Metacognition in mammalsand machines. Scientific American Mind 25(5), 35–35 (2014)
14. Gibney, E.: Google ai algorithm masters ancient game of go. Nature 529, 445–446 (2016)
15. Halliday, M.A.K., Matthiessen, C.M., Yang, X.: Construing experience through meaning: A language-based approach to cognition. MIT Press (1999)
16. Johnson, M.K.: A multiple-entry, modular memory system. Psychology of Learning and Motivation, vol. 17, pp. 81 – 123. Academic Press (1983)
17. Koch, C.: A theory of consciousness. Scientific American Mind 20(4), 16–19 (2009)
18. Koch, C., Tononi, G.: A test for consciousness. Scientific American 304(6), 44–47 (2011)
19. Markoff, J.: Computer wins on jeopardy!: trivial, its not. New York Times 16 (2011)
20. Mesnil, G., He, X., Deng, L., Bengio, Y.: Investigation of recurrent-neural-network architectures and learning methods for spoken language understanding. In: Interspeech. pp. 3771–3775 (2013)
21. Mikolov, T., Karafiát, M., Burget, L., Cernockỳ, J., Khudanpur, S.: Recurrent neural network based language model. In: Interspeech. vol. 2, p. 3 (2010)
22. Mira, J.J., Navarro, I., Botella, F., Borrás, F., Nuño-Solinís, R., Orozco, D., Iglesias-Alonso, F., Pérez-Pérez, P., Lorenzo, S., Toro, N.: A spanish pillbox app for elderly patients taking multiple medications: randomized controlled trial. Journal of medical Internet research 16(4), e99 (2014)
23. Oizumi, M., Albantakis, L., Tononi, G.: From the phenomenology to the mechanisms of consciousness: integrated information theory 3.0. PLoS Comput Biol 10(5), e1003588 (2014)
24. Overgaard, M., Overgaard, R.: Neural correlates of contents and levels of consciousness. Frontiers in psychology 1 (2010)
25. Pan, S.J., Yang, Q.: A survey on transfer learning. IEEE Trans. on Knowl. and Data Eng. 22(10), 1345–1359 (Oct 2010)
26. Picard, R.W.: Affective computing: from laughter to ieee. IEEE Transactions on Affective Computing 1(1), 11–17 (2010)
27. Picard, R.W., Picard, R.: Affective computing, vol. 252. MIT press Cambridge (1997)
28. Schmidhuber, J.: Deep learning in neural networks: An overview. Neural Networks 61, 85–117 (2015)
29. Sharkey, A., Sharkey, N.: Granny and the robots: ethical issues in robot care for the elderly. Ethics and Information Technology 14(1), 27–40 (2012)
30. Solomatine, D.P., Siek, M.B.: Modular learning models in forecasting natural phenomena. Neural networks 19(2), 215–224 (2006)
31. Tao, J., Tan, T.: Affective computing: A review. In: International Conference on Affective Computing and Intelligent Interaction. pp. 981–995. Springer (2005)
32. Tononi, G.: An information integration theory of consciousness. BMC neuroscience 5(1), 1 (2004)
33. Yamazaki, K., Watanabe, Y., Nagahama, K., Okada, K., Inaba, M.: Recognition and manipulation integration for a daily assistive robot working on kitchen environments. In: Robotics and Biomimetics (ROBIO), 2010 IEEE International Conference on. pp. 196–201. IEEE (2010)
34. Zhang, M.L., Zhou, Z.H.: A review on multi-label learning algorithms. IEEE transactions on knowledge and data engineering 26(8), 1819–1837 (2014)