CM3020 Artificial Intelligence

Name - Shivam Maheshwari
Student Number - 190349118

PART A - Reviewing research papers about robotic science

Introduction to the field of Robotic Science

What is Robotic Science?

Robotic Science is a way of automating scientific discoveries through the cycles of experimentation to test hypotheses and interpret findings without human guidance.[1]

What is it aiming to achieve?

Robotic Science is aiming to produce scientific discoveries by automating them without any human intervention. Automated scientific discoveries also help to minimise the cost to discover and develop along with saving time as the automated systems can work almost all the time, they don't need breaks, holidays or get tired like humans.

What kind of techniques does it use?

The techniques used in Robotic science are logical reasonings like Abduction, Induction and deduction. For eg. Robot Scientist Adam used abduction to generate hypotheses about gene functions. There are other techniques like Deep Learning, Neural networks and reinforcement learning which are also used in Modern automated systems for scientific discoveries and analysis.

Description of three papers about Robotic Science

1. **Cheaper faster drug development validated by the repositioning of drugs against neglected tropical disease. [1]**

- **What are the researchers trying to achieve?**

  Researchers are trying to make drug discovery cheaper and faster. The discovery of new drugs is exceptionally slow. There are many tropical and orphan diseases caused by parasites which kill millions of people each year but they are generally neglected due to economic reasons. So, the researchers demonstrate the viability of

the Robot Scientist ‘Eve’ in drug discovery. They focus on finding lead compounds for these tropical diseases.

- **Which techniques did they use to address the problem?**

  Researchers used Eve - Robot Scientist to generate key compounds need for drug development. Eve uses Active learning to select compounds to test its hypotheses along with reinforcement learning i.e. to find the compounds that have the best and optimal performance for the drug development.

  Eve’s modes including Library-screening, hit-confirmation and intelligent screening were used to demonstrate the automation of the drug development and generate the lead compound. The problem task was tightly defined by the human who engineered the assays. The human manual effort was also required to run and maintain ‘Eve’ and certain programs during the different stages of the cycle. Multiple iterations of active learning were conducted to find the cherry-picked compounds.

- **Which techniques did they use to evaluate their research?**

  Researchers used results from Eve’s active learning methods and compared it to the standard brute force way of conventional mass screening. The brute force way is very slow as each compound is manually tested. So, they decided to use empirical results from Eve against the Maybrdige HitFinder library which has over 14,400 compounds.

- **How successfully do you think they were in achieving their goals?**

  I think they did a great job in finding the compound for the treatment of the tropical disease caused by parasites. The results are quite favourable. The same Eve’s standardised assays and active learning can also be applied to the disease caused by bacteria, viruses etc. The drug development process can now not only be cheaper but also faster due to automated drug discovery.

- **Did they provide any source code or open data?**

  Yes, they did provide both the source code and dataset used for this research. The dataset is available at http://disc.brunel.ac.uk/eve-dataset. The code is available at https://github.com/RobotEve/ RobotEve which includes Eve low-level control, active learning software and other configuration files.

2. The Analysis of Yeast Cell Morphology Using a Robot Scientist [2]

- **What are the researchers trying to achieve?**

  Researchers are trying to analyse the morphology of the budding yeast cell by experimenting with four deletant strains. They use image analysis software of the Robot Scientist ‘Eve’ to automatically obtain quantitative morphology features of the yeast cell.

- **Which techniques did they use to address the problem?**

  Researchers used multiple techniques to address the problem as listed below:-

  1. Yeast cell segmentation - by computing the convolution of image matrices and Sobel filter.
  2. Calculation of shape feature of a yeast cell.
  3. Strain selection - selected 4 different strains which have different cell sizes and different shapes.
  4. Image acquisition - took images by calibrating the settings of yeast cell images.
  5. Parameter selection - They used the parameters such as Roundness, Bud neck position angle, and growth direction angle to describe the features of the yeast cell.

- **Which techniques did they use to evaluate their research?**

  The evaluation was done on the following parameters:-

  1. Roundness distribution of four strains - comparing roundness with other strains and matching the results. They found out that the roundness of one particular strain is significantly different from the rest three.
  2. The Comparison with SCMD - Researchers compared the results obtained from Robot Eve to the SCMD CalMorph software (an image processing program for analysing cell structure, nuclei, cell walls etc.). They found the results to be qualitatively different. Some strains tend to have a long axis of cells in SCMD results but have a small axis in the results obtained by their evaluation. The roundness in both the comparisons was also qualitatively different.

- **How successfully do you think they were in achieving their goals?**

  I think they were successful in analysing the Yeast cell morphology using a Robot scientist but there were discrepancies in the results obtained from their experiments and SCMD which they used as a benchmark for comparison. The results were quite

significantly different. They said it might be due to different growing environments or the technique to fix the cells before analysing in SCMD may have caused different results. I think more research and experiments are needed on yeast cell morphology.

- **Did they provide any source code or open data?**

  They haven't provided any source code in the research paper. They have mentioned the links for the SCMD software which they used for comparison. They have provided all the data obtained after the experiments along with the comparison data from SCMD in tabular form.

3. Automated Discovery in Chemistry Laboratory [3]

- **What are the researchers trying to achieve?**

  Researchers are trying to automate the discovery process in chemistry laboratories as opposed to human intervention. They use FAHRENHEIT, a discovery system to analyse quantitative regularities in electrochemistry experiments. The capabilities of FAHRENHEIT include finding empirical equations, discovering maxima locations, and heights and evaluating errors.
  Researchers selected electrochemistry for multiple reasons - the computer-operated sensors are readily available in the chemistry lab and it is relatively easy to automate various electrochemical operations on chemical samples.

- **Which techniques did they use to address the problem?**

  Researchers used FAHRENHEIT, an automated discovery system to report on the electrochemical operations. They used Cadmium, lead and copper ions in the +2 state for this experiment. Fahrenheit uses the following techniques to address the problem:-

  1. Finding points of special interest - such as maxima, minima, and discontinuities and use them in a recursive mechanism for the detection of multidimensional regularities.
  2. Determining the reproducibility of results - performing a number of experiments to analyse the data for regularities.
  3. Handling experimental errors.
  4. Establishing the communication between the FAHRENHEIT and PC including the command language used by FAHRENHEIT.

- **Which techniques did they use to evaluate their research?**

Researchers evaluated the FAHRENHEIT results according to accuracy, range of application and usefulness. For accuracy, they compared the findings with the accuracy achieved by human researchers. After comparing the results, they found out that the maxima and regularities are equivalent within empirical error.

They also found out that FAHRENHEIT can go far beyond electrochemistry and can deal with numerical data regardless of the application domain. Researchers also evaluated that the system can be improved a lot and the work done by a research assistant in a day is equivalent to 1 min of data analysis by FAHRENHEIT.

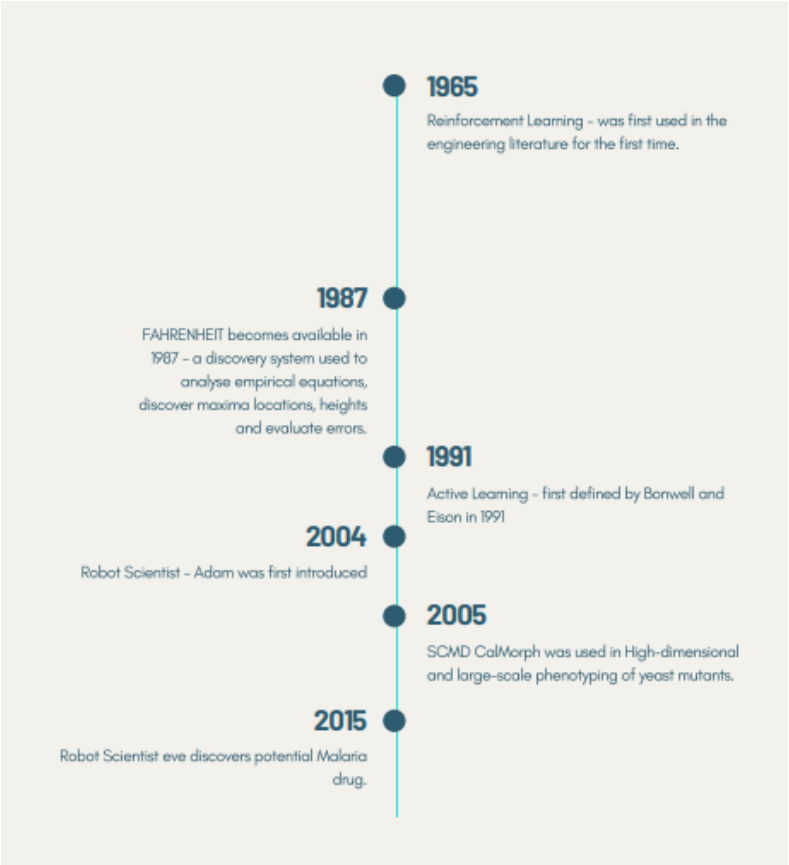
- **How successfully do you think they were in achieving their goals?**

  I think they were quite successful. They were not just able to automate the electrochemical experiments but also conclude that it would work in other domains as well which deal with numerical data. The scientist and ML researchers may save a good amount of time and effort on data analysis due to automation.

- **Did they provide any source code or open data?**

  They haven't provided any source code in the research paper. They have provided the data obtained from the experiments along with graphs to better understand the work.

Creating a timeline of developments in Robotic science



Discussion of ethics of Robotic science

1. Cheaper faster drug development is validated by the repositioning of drugs against neglected tropical diseases.

   Ethics are not explicitly discussed in the paper but the research itself is ethically correct and important for mankind. I think they are working for a very noble cause to discover the drug for neglected tropical diseases which has caused death to millions of people. They have followed ethical principles like Transparency (very transparent related to their analysis and findings, Openness (shared data regarding experiments and source code).

2. The Analysis of Yeast Cell Morphology Using a Robot Scientist

   Ethics are not explicitly discussed in the paper. They have been very transparent with their analysis and findings. They have also been very Honest about the discrepancy in the results obtained from their experiment compared to the SCMD CalMorph and concluded that more research is needed.

3. Automated Discovery in Chemistry Laboratory

Ethics are not explicitly discussed in the paper. They have provided all the experimental data along with the references they used for this research. They have also been very Honest and talked about the possibility of automating in other domains,s not just the electrochemical operations. They have also implied how a day's work of a researcher can be done in 1 min by the discovery tool.

Statement on the reliability of the references chosen

How reliable do you think the papers you read were?

The research papers are very reliable. All three papers which have been selected for this report are published in well-known journals and international scientific societies, organisations and government-approved sites like IEEE, AAAI and NCBI.

Do you think this work should be taken seriously? Why?

I think their work should be taken seriously. They have provided reports of deep analysis with multiple experiments to clearly demonstrate the benefits of using Robotic science in automating scientific discovery.

Reference List

[1] Kevin William, Elizabeth Bilsland, Andrew Sparkes, Wayne Aubrey, Michael Young, Larisa N Soldatova, Kurt De Grave, Jan Ramon, Michaela de Clare, Worachart Sirawaraporn, Stephen G Oliver, Ross D King “Cheaper faster drug development validated by the repositioning of drugs against neglected tropical diseases.” National Center of Biotechnology Information (NCBI), 2015.

[2] Yihui Liu, Katherine Martin, Andrew Sparkes, Ross D. King “The Analysis of Yeast Cell Morphology Using a Robot Scientist“. 2010 International Conference on Computational Intelligence and Security. IEEE, 2010.

[3] Jan M. Zytkow, Jieming Zhu, Abul Hussam “Automated Discovery in a Chemistry Laboratory”. Association for the Advancement of Artificial Intelligence, 1990.