University Teaching Qualification

Basiskwalificatie Onderwijs (BKO)

Teaching Portfolio

Dr. Pratik Gajane

13 September 2023

Chapter 0

Introduction

In this dossier, I shall endeavor to explain how I have pursued and will continue to pursue the following five competences.

1. Course design (see Chapter 1)
2. Teaching and supervising (see Chapter 2)
3. Assessment (see Chapter 3)
4. Evaluating teaching (see Chapter 4)
5. Professionalization (see Chapter 5)

I use 2AMM20 Research Topics in Data Mining : Reinforcement Learning Track as a basis for this dossier. I shall simply refer to it as the course, wherever it is clear from the context.

This course is offered by the Mathematics and Computer Science (M&CS) Department in the first quarter of an academic year. In the academic year 2022-2023, this course had the following three independent tracks.

- Exceptional Model Mining & Missing Data : Lecturers – Dr. Wouter Duivesteijn and Rianne Schouten.
- Anomaly Detection : Lecturer – Dr Stiven Schwanz Dias.
- Reinforcement Learning : Lecturer – me.

Dr. Wouter Duivesteijn was also the responsible lecturer for the whole course.

Other Teaching Experience

2022-23	Data Mining (Lecturer)	BSc	140 students	TU/e
2022-23	Embodying Intelligent Behavior in Social Context (Lecturer)	MSc	41 students	TU/e
2021-22	Data Intelligence (Project Supervision)	MSc	50 students	TUe
2013-14	Data Mining (TA)	BSc	20 students	IIT Madras
2013-14	Introduction to Machine Learning (TA)	BSc	60 students	IIT Madras
2012-13	Computational Engineering (TA)	BSc	50 students	IIT Madras
2012-13	Introduction to Research (TA)	BSc	100 students	IIT Madras

Table 1: Other Teaching Experience

Chapter 1

Course Design

1.1 General Introduction – 2AMM20 Research Topics in Data Mining : Reinforcement Learning Track

I undertake the design of the Reinforcement Learning (RL) track in the course 2AMM20 Research Topics in Data Mining. This course is an elective course worth 5 ECTS credits.

1.2 Entry Level of the Students

This course is taken mostly by the students in the first year of Master of Data Science and Artificial Intelligence (DS&AI) curriculum. The following are the prerequisites for the course.

- Elementary statistics and probability theory.
- Comfort with applying mathematical tools.
- Bachelor's course worth of background knowledge in Data Mining and Machine Learning.

1.3 Designing the Course Matching the Vision of TU/e, the Department and My Personal Vision on Education

In section 1.3.1, I outline the three visions and the commonalities between them. In Section 1.3.2, I describe how I designed the course accordingly.

1.3.1 TU/e Vision, Departmental Vision and My Personal Vision on Education

The TU/e vision on education has set out the following five major goals :

1. Educate engineers for the future.
2. Serve diverse learners.
3. Create personal learning paths.
4. Transform from teaching to learning.
5. Offer challenge-based learning (CBL).

On speaking with the responsible lecturer for the course, Dr. Wouter Duivesteijn, I understood the following to be the major part of the vision of M&CS Department for this course.

- The course should be advanced enough so that on successful completion of the course, students are able to produce novel research.

My own vision of teaching is outlined in Section 2.1. A common theme among these three visions could be described as creating an accessible course based on the paradigm of CBL that is suitable for a variety of learners and that will help the students to produce novel research.

1.3.2 Course Design according to the Vision

According to the vision described in the last paragraph, firstly the Reinforcement Learning Track of the course was designed to be a CBL course. The timeline was as follows :

- In the first plenary lecture, Dr. Wouter Duivesteijn would introduce the three tracks, the respective instructors, the goal and setup of the course, expectations from the students, and lecturer support. In particular, he explained how the students can influence which track they will be assigned to. As the students could influence the track assignment, I believe it helped the students to take ownership of their learning which is in line with the fourth goal outlined in the TU/e vision in Section 1.3.1.
- In the next five plenary lectures over 2.5 weeks, I would introduce various topics in reinforcement learning and the state-of-the-art. The focus of the lectures would be understanding reinforcement learning via mathematical analysis. I believe this would serve as a complementary learning experience to other courses such as 2AMC15-Data Intelligence Challenge. I was involved in the course 2AMC15 and my conversations with students revealed that some of them would like to understand the mathematical foundations of reinforcement learning. Accordingly, I designed the plenary lectures with the following Content Tree as given in Fig 1.1. The

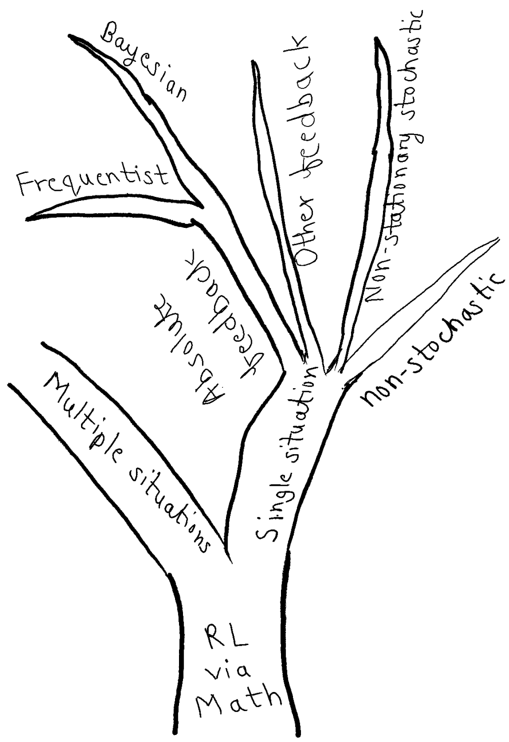


Figure 1.1: Content Tree for the Course

Content Tree depicts the major concepts in the course plan. Creating this Content Tree helped me to select the topics and organize them in the lectures as given in Table 1.1.

I create lecture slides taking Meyer's principles of media design [1] into consideration. For example, I try to eliminate extraneous material and focus on relevant content to minimize distractions and improve learning efficiency. I highlight important information using cues, such as arrows, bold text, or highlighting, to guide learners' attention and emphasize key concepts. I place related text and images close together on the screen to help learners make connections more easily and improve comprehension. I present corresponding text and images simultaneously, rather than sequentially, to facilitate understanding. I break content into smaller, manageable segments to help learners process and understand the material more effectively. I provide students with an overview of the content and its structure before diving into the details, enabling them to build a framework for organizing new information.

In the lectures, I try to incorporate different forms of learning to accommodate a variety of learners. The lecture slides are created using a

Lecture	Topics
1	Introduction to Reinforcement Learning
2	A frequentist algorithm for single-state RL with absolute feedback
3	A Bayesian algorithm for single-state RL with absolute feedback
4	Algorithms for other feedback and non-stationary stochastic environment
5	An algorithm for multiple-state reinforcement learning

Table 1.1: Lectures Topics

consolidated color scheme and they contain illustrative figures to benefit visual learners. For auditory learners, I plan my lectures with scheduled questions and invite students to answer them. For hands-on learners, I recommend trying the mathematical analysis taught in the lecture on their own. If they do not understand any aspects of the mathematical analysis, I answer them promptly and also provide them positive feedback on the parts they managed to complete correctly on their own. See an example of the above in the email conversation given in Appendix F. In the email conversation, student identification details are redacted.

I shared some of the lecture slides with Dr. Raoul Nuijten (Lecturer, Department of Industrial Engineering and Innovation Sciences) and incorporated his feedback.

- In the last 5 weeks of the course, the students will work on their research project with the stated aim of producing novel research. Keeping in line with the focus of the course, students will be asked to choose a RL problem, devise a solution for it, and analyze it mathematically. During these 5 weeks, I conduct regular meetings with students and guide them on their research projects. Via these meetings, I can personalize supervision for students and guide them toward creating novel research.

1.4 Connection to Research and Future Fields of Occupation

This is a research-oriented course as we will shortly see in the following sections. This course can help the students to start their career in research laboratories or pursue further education.

1.5 Learning Objectives of the Course

When this course was offered previously (2022-23 GS1), I had set the following learning objectives.

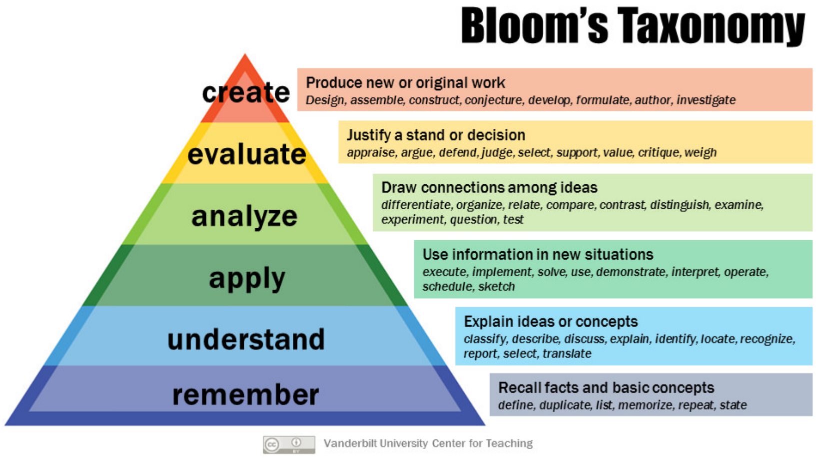


Figure 1.2: Blooom's Taxonomy

1. To gain an understanding of various reinforcement learning problems and formulate them mathematically.
2. To devise solution strategies for these problems.
3. To prove mathematical performance guarantees for these solutions.

On reflection, I have decided to set the following learning objectives for the next iteration of the course.

1. You will be able to formulate reinforcement learning problems mathematically.
2. You will be able to construct solution strategies for reinforcement learning problems.
3. You will be able to prove mathematical performance guarantees for the constructed solution strategies.

Note the following improvements made in the learning objectives.

- More measurable: "Gain an understanding" → "formulate". The latter is measurable unlike the former.
- Use of the phrase "You will be able to…": This was done following the recommendation from a UTQ module.

These learning objectives cover the levels of understanding, applying, analyzing, evaluating and creating from Bloom's Taxonomy given in Fig 1.2. The level of learning outcomes is appropriate for this master's level course.

1.6 Connection to the Learning Outcomes of the Degree Program

The learning outcomes of the DS&AI master's degree program are as follows:

1. Knowledge and understanding in DS&AI.
2. Applying knowledge and understanding in DS&AI.
3. Making judgments and proficiencies in research and design in DS&AI.
4. Communication.
5. Learning skills and attitude.

The learning outcomes of the degree program are connected to the learning objectives of the course (see Section 1.5). In particular, the first learning objective of the course corresponds to the first and the second learning outcome from the above list. The second learning objective of the course corresponds to the second learning outcome from the above list. The third learning objective of the course corresponds to the third learning outcome from the above list. The fourth learning outcome from the above list is covered by asking the students to write a detailed report on their research project (to be explained shortly in Section 1.8). And finally, since this is a CBL course, students are rewarded by being independent, motivated, and self-actualized self-learners which enables them to fulfill the fifth learning outcome from the above list.

1.7 Learning Activities

The learning activities for the students in the 2022 iteration of the course were as follows:

- Attend lectures. – The education students receive via these lectures is directly connected to the ongoing research in the field of reinforcement learning. The lectures cover a broad array of research topics in the field (see Content Tree given in Fig 1.1). In each of the lectures, I introduce various research problems and teach the students how to formulate them mathematically. Then I show how to construct various solutions for the introduced problems. Here, I first teach them some naive strategies and explain them how they (mostly) fail to solve the problem at hand. Then I gradually lead them to suitable solutions. This helps the students to develop the ability to construct solutions incrementally, which I believe is a key research skill. I also teach them various proof techniques which will be useful for them in proving the associated mathematical guarantees.

- Read books and papers listed on Canvas for guided self-study – The lectures are supported by books and papers listed on Canvas and I include links to the relevant additional study material in the slides.

- Guided online discussions – In Review Phase (see Fig 1.3), students are asked to discuss the research papers online mimicking a real conference review system. This digitally enhanced learning activity provides the students with an asynchronous mode of learning. At times, I guide these discussions to make sure that the significant points are covered. These discussions help the students to appreciate how research problems are formulated and solved.

In the above, it can be seen that the learning activities are purposefully designed to help the students to achieve the learning objectives given in Section 1.5. On reflection and following students' evaluations, I have decided to add a few more learning activities as detailed in Section 1.10.

1.8 Assessment Plan

In lieu of the departmental vision, the end goal of this course is that students will make a research contribution: they will go beyond the current state of the art. Therefore, this course is designed according to the paradigm of Challenge-Based-Learning and the students had to complete three assignments as detailed in Figure 1.3.

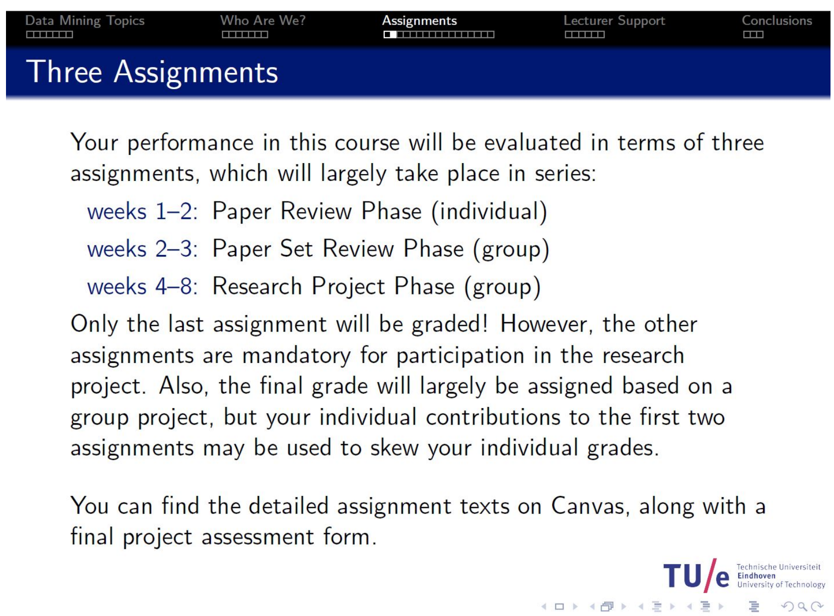


Figure 1.3: Assignments in the course 2AMM20

1.8.1 Summative Assessment

Students are asked to form a group of five of their own choosing. During weeks 4-8, each group works on a research project in reinforcement learning. The groups have the freedom to choose the topic of their respective research projects.

The grade received by a student taking this course is fully determined by their research project. In the Research Project Phase (i.e., weeks 4-8), each group is entitled to a weekly meeting with the lecturer. For their project, each group has to submit a detailed report which forms the basis of the assessment.

1.8.2 Formative Assessment

In weeks 1-3, students work on Assignment 1 and Assignment 2. In these assignments, students are asked to review published scientific articles mimicking the real-world reviewing process. The lecturer can view the ensuing online discussion and may choose to participate in it to foster the discussion further.

As for the weekly meetings, each group is asked to send a brief progress report consisting of

- summary of their work on the project till then (focusing on the work since the preceding weekly meeting),
- description of their ideas that worked (and also those which did not work), and
- their plan for the subsequent week.

These reports form the basis of the discussion during the weekly meetings. Moreover, they also inform the lecturer of the ongoing progress of each group.

1.9 Evaluation Plan

I made use of the following methods for evaluating this course.

- Peer-evaluation : While I was designing the course and the lecture materials, I gave 2 demo lectures for my peer coach. I incorporated his feedback to make certain changes to the lectures and to improve them further.
- Self-evaluation : All the lectures in 2022 iteration of the course were recorded. After each lecture, I used to watch the corresponding recording to identify the positives and the negatives. These recording also helped me to keep track of the questions asked by students. In particular, on reviewing a lecturer, I realized that my answer to a question asked by a student could be made more precise. In the subsequent lecture, I recalled that question and provided the students with the precise answer.
- Informal discussion with students.
- Official students' evaluation (enclosed in Appendix A).

- Online questionnaire (enclosed in Appendix B) : I also conducted an anonymous online survey amongst students to receive feedback on issues not covered in the official students' evaluation.

1.10 Reflections

What worked well :

1. Design of course materials. See Figure B.1 in Appendix B.
2. Inclusive and diverse illustrative examples. See Figure B.2 in Appendix B.

On the other hand, I identified the following areas of improvement.

1. Some students were not able to do the mathematical analysis as part of their research project. On speaking with some of these students, I realized that most of them possessed the skills to do the mathematical analysis but lacked the experience.
2. A few parts of my lectures were not in accordance with some of Mayer's principles of multimedia design [1] e.g., the Redundancy Principle and Temporal Contiguity Principle. In violation of the Redundancy Principle, some of the figures in the lecture slides could have been distracting. And Breaching the Temporal Contiguity Principle, sometimes I presented the corresponding narration and slides successively.

Based on these reflections, I have decided to add the following learning activities for the next iteration of the course :

1. An optional online pre-test before the course for the students to assess their prerequisite skills and knowledge for the course. After completing the test, the students will be given model solutions and further reading materials related to each question. Thus, students will be able to review and possibly refine the skills and knowledge which will help them throughout the course.
2. Guided in-class construction of proofs with feedback from the lecturer. This will provide the students the experience to do the mathematical analysis independently as well as collaboratively.

I am also going to adapt my teaching style to incorporate Redundancy Principle and Temporal Contiguity Principle from Mayer's principles of multimedia design [1].

I presented the above improvement plan to my fellow teachers in the UTQ module titled Designing Courses & Projects. Their feedback can be seen in Figure 1.4 and Figure 1.5.

A noteworthy tip given in this feedback is that students might find it intimidating to do the proofs on their own. This was for my initial idea which was to

Figure 1.4: Feedback from the UTQ module – Designing Courses & Projects

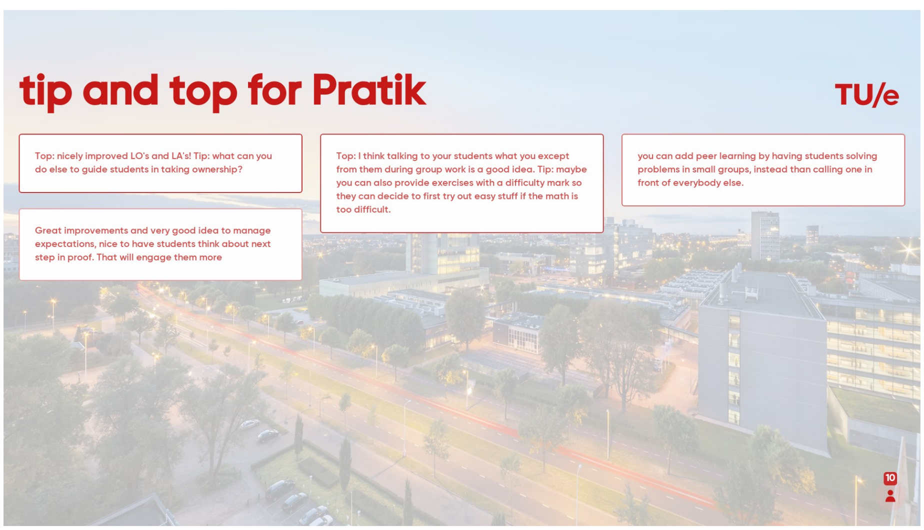


Figure 1.5: Feedback from the UTQ module – Designing Courses & Projects

invite students to do the proof on the board during the class. Another related tip is to add peer learning by having students do the mathematical analysis in small groups. Following these tips, I have amended my initial idea to asking students to construct the mathematical analysis in small groups. Another useful tip that I plan to incorporate is to provide the difficulty level of each in-class exercise and problems in the pre-test. Using these provided difficulty levels, students can first attempt easier problems before moving on to more difficult problems.

Supporting Products

- Official students' evaluation enclosed in Appendix A.
- Online questionnaire enclosed in Appendix B.

Chapter 2

Teaching and Supervising

2.1 Teaching Philosophy

My teaching philosophy is built primarily on the following two pillars :

- Facilitating learning catered to students' background and their academic level – For example, if I am teaching machine learning to second-year Computer Science master's students, I would incorporate strategies from the challenge-based learning paradigm via first teaching them the Big Idea, then assisting them to investigate and formulate a challenge and finally guiding them toward scientific solutions to the identified challenge. On the other hand, if I am teaching machine learning to second-year Industrial Design bachelor's students, I would use more conventional instructional teaching with regular assignments to develop their basic understanding of the topic.

- Inclusive and accessible education – In order to provide diverse and inclusive teaching and supervision, I act according to the following guidelines :

Principle	Activities
Positive class climate	Learning names, in-class surveys and activities
Explicit expectations	Clear assessment criteria, timely feedback
Diverse course content	Use of multiple and diverse examples
Accessible course	Use of dyslexia-friendly fonts (e.g. Arial)
Commitment to inclusion	Self-inventory of biases, ways to overcome them

Table 2.1: Guidelines for inclusive and accessible education

2.2 Preparing an Educational Meeting

I taught the Reinforcement Learning (RL) track in the course 2AMM20 Research Topics in Data Mining in 2022-23 GS1. In this section, I will focus on Lecture 3 of this track.

2.2.1 Learning Objectives

The learning objectives of Lecture 3 were as follows :

- LO1 : Students will be able to recall and explain the problem setting of Stationary Stochastic Bandits and the frequentist perspective of solving it.
- LO2 : Students will be able to solve Stationary Stochastic Bandits from a Bayesian Perspective.
- LO3 : Students will be able to explain a Bayesian solution, called Thompson sampling, for the above problem.

2.2.2 Preparatory assignment

Students were asked to watch the video lecture for the previous lecture and revise the core concepts, especially if they were absent from the class.

2.2.3 Materials for the meeting

I prepared lecture slides making use of colors, overlays, figures and graphs to assist the students in understanding the topics being taught. I explain the purpose for their use below.

- Colors : This lecture is about devising an algorithm and analyzing it mathematically so naturally the lecture slides involve plenty of mathematical concepts and notations. To aid understanding, I use consistent colors for concepts/notations throughout the lecture slides (in fact, throughout the course) e.g. green for rewards, blue for policies, red for regret etc. The goal of such a color scheme is that students will be able to understand these concepts/notations quickly using a visual mode and link them to other similar concepts/notations.
- Overlays : With the help of overlays, I display the relevant content on slides when I am speaking about it during the lecture. Firstly, this is aligned with Mayer's principles of multimedia design [1]. Secondly, I believe this avoids overwhelming/distracting the students with lots of content on the slide when I am not speaking about the concerned topic.
- Figures and graphs : The purpose of using figures and graphs is that some concepts like intervals (see slide 5 in the lecture slides given in Appendix H), probability distributions (see slide 8 in the lecture slides given in Appendix H) are easy to understand via a visual mode.

2.2.4 Lesson plan

The following was the Lesson Plan for the lecture I delivered on September 14, 2022. Note that the Lesson Plan was prepared (and the lecture was delivered)

LO	Time	Activity
LO1	2-3 minutes	A quick recap and reintroducing Stationary Stochastic Bandits
LO1	10 minutes	Viewing the algorithm taught in Lecture 2 from a frequentist perspective
LO2	5-6 minutes	Introduction to Bayesian perspective for solving Stationary Stochastic Bandits
LO2	2-3 minutes	Introduction to a concept required to construct the algorithm (Beta distribution)
LO2	3-4 minutes	Elaborating on the Bayesian perspective for solving Stationary Stochastic Bandits
LO3	5-6 minutes	Introduction to Thompson Sampling
LO3	15-17 minutes	Explaining how Thompson Sampling works

Table 2.2: Lesson Plan

prior to me attending any of the UTQ modules. After attending the UTQ modules and reflecting on the lecture, I have come to realize that the lecture could be improved particularly with respect to student participation. The Improved Lesson Plan, given in Table 2.3, allows for more student participation and takes into consideration other feedback given by my peer coach, Dr. Wouter Duivesteijn (see Appendix C), my UTQ coach, Ms. Hester Morssink (see Appendix D) and the students from RL track (see Appendix A and Appendix B).

Mentimeter Quizzes 1-5 mentioned in the Improved Lesson Plan are shown in Figures 2.1, 2.2, 2.3, 2.4, and 2.5 respectively. The time steps mentioned in Figures 2.3, 2.4 and 2.5 are explained in slides 15, 16 and 17 respectively in the lecture slides enclosed in Appendix H.

2.3 Conducting an Educational Meeting and Reflection on the Performance

In this section, I shall provide a detailed overview of the way I taught Lecture 3 using the Lesson Plan given in Table 2.2. I will interleave the overview with my reflections and related points mentioned in the feedback given by Dr. Duivesteijn, Ms. Morssink and the students from RL track. Furthermore, I shall explain how I have incorporated findings from these reflections and the feedback in the Improved Lesson Plan given in Table 2.3.

As seen in Table 2.2, I started Lecture 3 (See slide 2 in Appendix H) with a quick recap and a reintroduction of Stationary Stochastic Bandits. The problem setting of Stationary Stochastic Bandits was studied extensively in the previous

LO	Time	Teacher Activity	Student Activity
LO1	4 minutes	Inviting students to think-pair-share on core concepts from previous lectures	Practise recalling and explaining core concepts e.g. Stationary Stochastic Bandits
LO1	1 minute	Recap of previous lectures	Recall previously taught concepts
LO1	2 minutes	Mentimeter Quiz 1	Recall Stationary Stochastic Bandits
LO1	2 minutes	Brief reintroduction to Stationary Stochastic Bandits	Consolidate understanding of Stationary Stochastic Bandits
LO1	6 minutes	Viewing the algorithm taught in Lecture 2 from a frequentist perspective	Recognize the frequentist perspective
-	1 minute	Outlining the structure of the lecture	-
LO2	4 minutes	Introduction to Bayesian perspective for solving Stationary Stochastic Bandits	Recognize the Bayesian perspective
LO2	2 minutes	Introducing a concept required in Bayesian algorithm (Beta distribution)	Identify Beta distribution
LO2	2 minutes	Mentimeter Quiz 2	Recognize differences between the frequentist and the Bayesian perspective
LO2	3 minutes	Elaborating on the Bayesian perspective	Consolidate understanding of the Bayesian perspective
LO3	4 minutes	Introduction to Thompson Sampling	Identifying construction of Thompson Sampling
LO3	14 minutes	Explaining how Thompson Sampling works using Mentimeter Quiz 3, 4 & 5	Discover the modus operandi of Thompson Sampling

Table 2.3: Improved Lesson Plan

lecture along with a solution for this problem. Lecture 3 also focuses on the same problem and constructing another solution for it. And that is why I decided to start Lecture 3 with a recap of Stationary Stochastic Bandits.

Improvement following feedback and reflection: While reviewing the video-recording of the lecture, I realized that LO1 (see Section 2.2.1) will be better served if students themselves recap the problem setting of Stationary

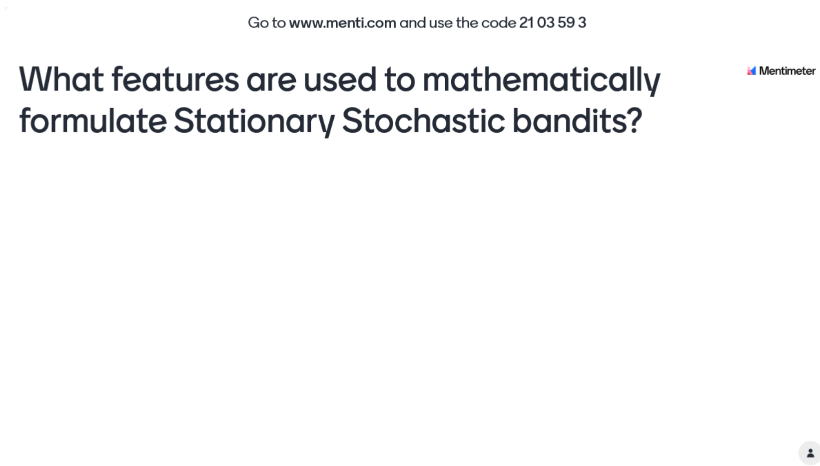


Figure 2.1: Mentimeter Quiz 1

Figure 2.2: Mentimeter Quiz 2

Stochastic Bandits. This was also mentioned by Ms. Morssink in her feedback (see the section of *Observed teacher behavior* in Appendix D). Therefore, I have introduced a think-pair-share activity at the beginning of the Improved Lecture Plan given in table 2.3. Here, I would invite the students to think-pair-share [2]. I would ask the students to recall the core concepts from previous lectures (including the problem settings) individually for a minute. Then, I would ask them to discuss their recollections in pairs for another minute. For the next two minutes, I would invite volunteers to share their recap with the class. Beyond achieving LO1, another objective of this activity is to increase student partici-

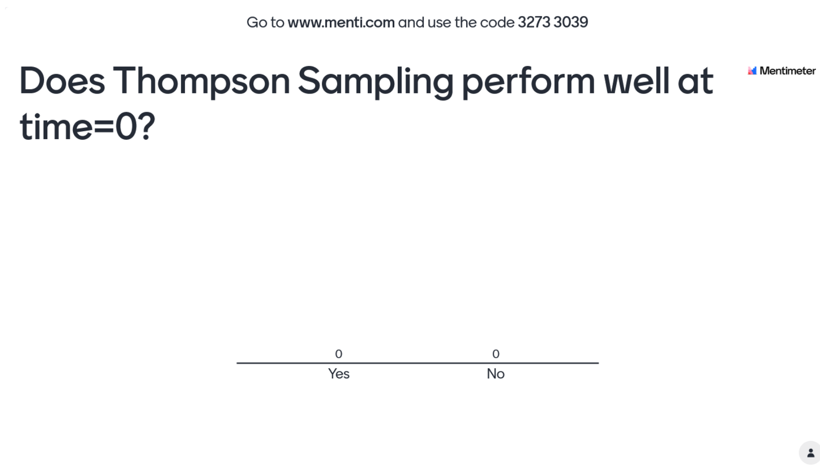


Figure 2.3: Mentimeter Quiz3

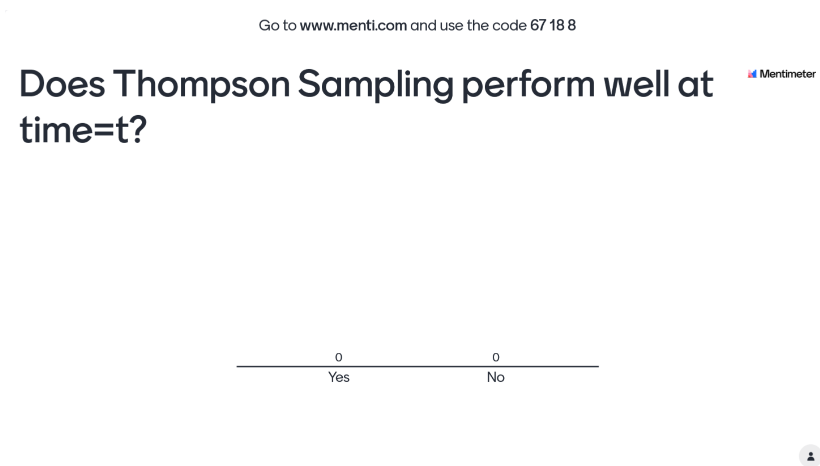


Figure 2.4: Mentimeter Quiz 4

pation [1]. This activity would also inform me if students have completed their preparatory assignment (see Section 2.2.2).

Next, I would provide a brief recap of Stationary Stochastic bandits. This would serve to fill the gaps in the understanding as indicated by the above think-pair-share activity.

Improvement following feedback and reflection: Here I would ask students to answer the question posed by Mentimeter Quiz 1 (see Figure 2.1).

[1] In a study across 4 German secondary schools, Mundelsee and Jurkowski [3] have indicated that think-pair-share increased the likelihood of student participation (measured using the proxy of hand raising).

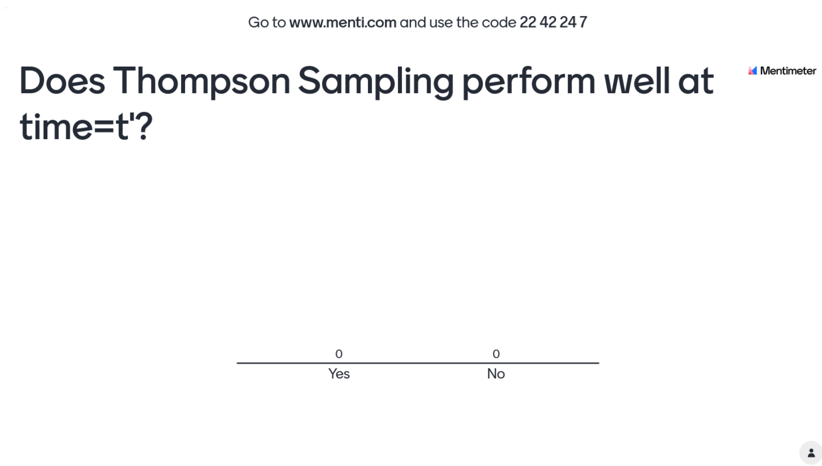


Figure 2.5: Mentimeter Quiz 5

The results will be shown to the class using a word cloud. This would further enhance student participation.

The next activity would be to view the algorithm taught in the previous lecture from the point of view of a frequentist perspective. This would help the students to recognize and recall the frequentist perspective of solving Stationary Stochastic Bandits (part of LO1 given in Section 2.2.1).

At this point, Lecture 3 would move on to the main objective of the lecture i.e., construction of a Bayesian algorithm for Stationary Stochastic Bandits.
Improvement following feedback and reflections: As mentioned by Dr. Duivesteijn (see the section of *Body* in Appendix C), even though the lecture reflected a clear structure, it was not clearly apparent to the students. Thus, at this point in the lecture, I would outline the structure of the lecture.

Next, I would introduce the Bayesian perspective of solving Stationary Stochastic Bandits (see slide 7 in the lecture slides in Appendix H). Following that I would introduce Beta distribution (see slide 8 in the lecture slides in Appendix H) which is a concept required in the construction of the Bayesian algorithm which we would see next.
Improvement following feedback and reflection: At this point, I would ask the students to answer the question posed by Mentimeter Quiz 2 (see Figure 2.2). The students will be shown two solutions and asked to match these solutions with either Bayesian or frequentist technique.

Next, I would further elaborate on the Bayesian perspective of solving Stationary Stochastic Bandits (see slide 9 in the lecture slides in Appendix H). I use this technique of cycling around the main concept repeatedly as this was one of the suggestions of Winston [4] and in my experience, it is an effective way of teaching.

Next, I would introduce a Bayesian algorithm called Thompson sampling (see slide 9 in the lecture slides in Appendix H). Note the use of overlays and colors on the slide. As discussed in Section 2.2.3, overlays help the students to focus on the particular point being taught and colors help the students to link similar concepts to each other. Finally, I would explain to students how Thompson sampling works (slide 11 onward in the lecture slides in Appendix H) and we would start building toward the mathematical guarantees provided by Thompson sampling. Here, I make an explicit callback to a similar-looking algorithm taught in the previous lecture (see at 35:35 in the enclosed video lecture) and explain to the students what differentiates Thompson sampling from that algorithm. In my experience, understanding the differences and similarities between two algorithms paves way for a clearer comprehension of both algorithms. **Improvement following feedback and reflection:** After displaying slides 15, 16 and 17 from the lecture slides in Appendix H), I plan to show the questions shown in Figure 2.3, 2.4 and 2.5 respectively. These questions would give the students some opportunities to reflect upon Thompson sampling which will help them to explain how the algorithm works (LO3 given in Section 2.2.1).

2.4 Reflections on My Teaching

What worked well:

1. Accessible learning environment – See Figure B.2 in Appendix B.
2. Explaining difficult topics in a clear and comprehensive manner – See the official students' evaluation in Appendix A and also the feedback I received via a student via Email in Appendix E.
3. Approachable and passionate teaching – See feedback from my peer coach Dr. Wouter Duivesteijn in Appendix C.

On the other hand, I identified the following areas of improvement.

1. Effective and clear speaking: Sometimes I tend to speak too fast and the pronunciation of certain words/phrases is hurried. This has been a recurring issue for me over a number of years and I have been working on it. To counter this tendency, I practice my lectures to gauge the right speed of narration. Although this has helped me to alleviate this problem to a certain extent. This can be seen in the feedback given by Dr. Wouter Duivesteijn in Appendix C.
2. Handling of disruptive students: I think I struggle with handling disruptive students. This not only breaks my flow during teaching but can also be bothersome to other students.
3. Presence in the class: During my previous teaching experience, I noticed that I used to be very static in the class. Thus while delivering lectures now, I tend to move around the classroom. I believe this helps in making the speaker more approachable.

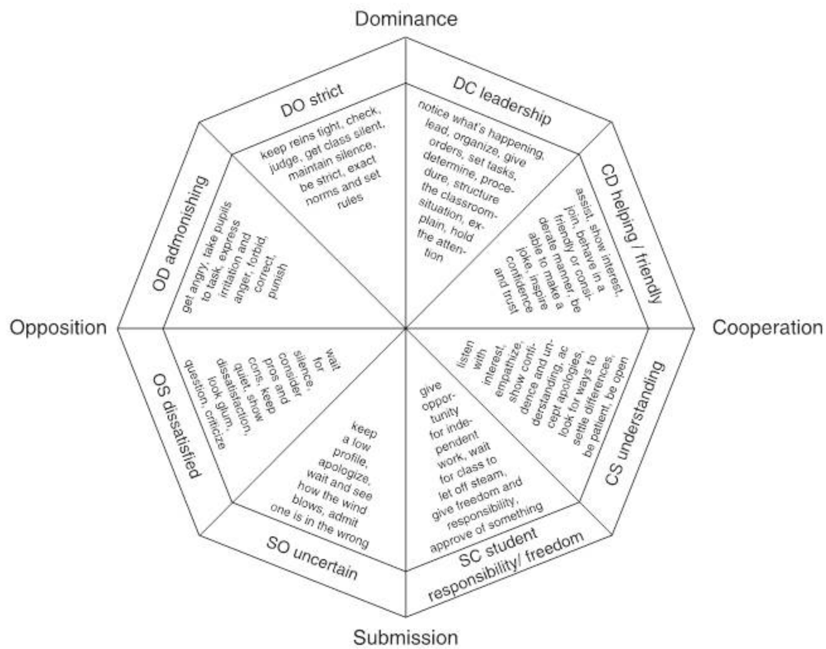


Figure 2.6: Model for Interpersonal Teacher Behavior

2.5 Feedback on Demo Lecture in the UTQ Module on Teaching Skills

I gave a demo lecture to my fellow teachers and the instructor in the UTQ module on teaching skills and I asked them to provide me feedback, particularly on the above three points. As for effective and clear speaking, I received feedback from the instructor that the narration is mostly good although some phrases are still hurried. So I am going to continue working on this issue. I have also enrolled in a TEACH module on powerful communication and voice techniques which might be helpful. As for the presence in the class, many of my fellow teachers and the instructor in the UTQ Module mentioned that I walk around too much while delivering the lectures and it can be distracting. Considering this feedback, I am going to cut down on the movement around the class during the lecture.

For handling the disruptive students in the class, I decided to make use of the Model for Interpersonal Teacher Behavior [5] which is reproduced in Figure 2.6. Wubbels et al. [5] categorize teacher behavior into four quadrants of the wheel shown in Figure 2.6. Furthermore, they argue that when a teacher's behavior falls into one of the quadrants, the students tend to display behavior from the diametrically opposite quadrant. Thus to deal with disruptive students

PhD		
2022-2023	Vishnu Veparala	Continual Learning
2021-Present	Danil Provodin	Constrained Sequential Learning (in collaboration with KPN)

MSc		
2022-2023	Ricardo v. d. Aa	Predictive Models for Inventory Control (in collaboration with Optiply)
2022-Present	Joost v.d. Haar	Supply Chain Management using ML (in collaboration with ASML)
2022-Present	Jiong Li	Exploration in Reinforcement Learning with Sparse Rewards
2022-2023	Wouter v. d. Wee	Curiosity-driven Fairness in Reinforcement Learning

Table 2.4: List of Supervised Students

Course	Number of groups	Objective
2AMM20	7 (each of 5 students)	Course project for the track of Reinforcement Learning
2AMC15	10 (each of 5 students)	Course project for Data Intelligence Challenge
DBM140	1 (5 students)	Extension of course project into a paper

Table 2.5: List of Supervised Groups

(Opposition quadrant), I decided to act in a stricter manner. I also plan to give a warning during the first lecture of the course that disruptive behavior will lead to penalties. I tried this warning during the demo lecture in the UTQ module on teaching skills. While most of the attendees said that the warning could help with preventing disruptive behavior, some attendees also said that the warning could be given in a more subtle and friendlier manner. Taking this feedback into consideration, I am going to continue working on how to handle disruptive students.

2.6 Supervising Students Individually and in Groups

Currently, I am a co-supervisor for 2 PhD and 4 MSc students with details given in Table 2.4. Additionally, I have supervised 18 groups with details given in Table 2.5.

2.6.1 Vision for supervision

My vision for supervising students consists of the following points :

- Enhance their problem-solving skills – Throughout the supervision period, I try to engage and enhance students' problem-solving skills via following the paradigm of challenge-based learning. This is aligned with the TU/e strategy 2030.

- Extend the state of the art – I set the goal of extending the state of the art for all the students that I supervise. This goal has already been successfully met by some of them and we have published the following papers – Li and Gajane [6], van den Broek et al. [7], Provodin et al. [8], van Tuijn et al. [9].

2.6.2 Method of supervision

During my first meeting with students/groups I supervise, I set clear expectations – I tell them what they can expect from me and what I expect from them. I also seek to recognize their own expectations of my supervision and of the project/research. By the end of the first meeting, we arrive at a clear understanding of these expectations.

Thereon, I mostly prefer to follow the collaborative supervision style wherein students and I both drive the work forward. However, I realize that some students at times require clear direction so I adapt my supervision style according to the situation and the student's needs. This was the case with one of my PhD students recently and I was able to assist him in submitting his first paper to a major conference. Here, I set a clear objective for him (submit a paper to the conference AAAI 2023) well in advance, conducted regular meetings with him, and worked with him extensively in the period prior to submitting the paper.

2.6.3 Setting up timeline

This varies depending upon the period of supervision. For course projects, the period of supervision is usually a couple of weeks so the timeline is rigorous with routine checks to see if expected progress is being made. On the other hand, for PhD students, the period of supervision is 4-5 years so here I set up a flexible timeline with them which is adaptable to their needs and the situation.

2.6.4 Feedback

I ask the students I supervise to send me regular progress reports and provide them with detailed written feedback with actionable points of improvement. I make sure to provide them with positive reinforcement to motivate them further.

2.6.5 Development of academic skills

I support students in their development of academic skills. For example, on realizing that one of my MSc students needed to improve her writing skills, I gave her some tips from my own experience and directed her toward other resources like TU/e career academy and an online course on scientific writing.

2.7 Reflections on My Supervision

What worked well:

- Positive reinforcement – I believe that my style of giving students positive reinforcement motivates them to perform better.
- Development of academic skills – I help students develop academic skills like academic writing, publishing, and creating presentations.

See the feedback on my supervision given by a PhD student in Appendix G.1 and by a master student in Appendix G.2.

On the other hand, I think I could improve my supervision of students who do not seem to produce considerable output. I dealt with this issue with one of the PhD students. I had a number of conversations with them where I tried to understand the possible reasons behind their lack of output. Furthermore, I set up a timeline for them to provide them with a structured environment. This worked as a short-term solution as the student was able to produce some output over the next couple of months. I would like to continue to explore other successful ways to deal with such students.

Supporting Products

- Official student evaluation for the course 2AMM20 in Appendix A
- Results from student survey from the RL track in the course 2AMM20 in Appendix B.
- Feedback on teaching activity filled by Dr. Wouter Duivesteijn in Appendix C.
- Feedback on teaching activity filled by Ms. Hester Morssink in Appendix D.
- Feedback on teaching by a student in Appendix E.
- Feedback on supervision given by a PhD student in Appendix G.1.
- Feedback on supervision given by a master student in Appendix G.2.

Chapter 3

Assessment

As a basis for this chapter, I shall use the CBL group project assignment in the RL track of the course 2AMM20.

3.1 Assessment Design

In this section, I describe the assessment for this course. I consulted with the responsible lecturer for the course, Dr. Wouter Duivesteijn, and this assessment does not violate any rules and regulations set by the TU/e.

- **Summative Assessment**
  For this course, the CBL group project assignment is the only summative assessment.

  In lieu of the departmental vision, the end goal of this course is that students will make a research contribution : they will go beyond the current state of the art and produce their own novel research. This is reflected in the learning objectives as well (see LO2 and LO3 in Section 1.5). Hence, a challenge-based project is a suitable assessment method as the students will get to experience the typical research life-cycle while working on their project.

  The reason for it being a group-based project rather than an individual project is two-fold. Firstly, workload management for students. If this were an individual assignment, the workload for students would have been too high, especially considering that it is a quartile-long course rather than a semester-long course. And secondly, research works are usually collaborative efforts. So working on a CBL group project, students will be able to hone their collaboration skills in order to reach a common research objective.

  Students are asked to form a group of five of their own choosing. During weeks 4-8, each group works on a research project in reinforcement learning. The groups have the freedom to choose the topic of their respective

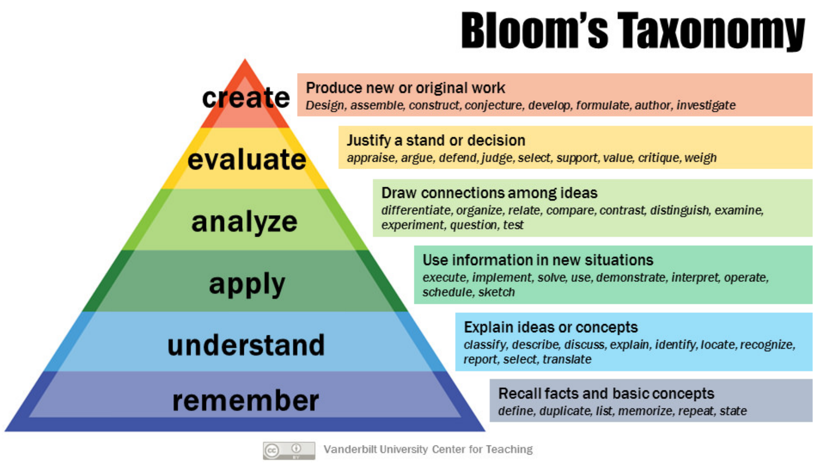


Figure 3.1: Blooom's Taxonomy

Assessment Matrix		
Goal 1	Goal	Formulate reinforcement learning problems mathematically.
	How tested	CBL – group project
	Level	Understand and apply
	Weight	33.3333…%
Goal 2	Goal	construct solution strategies for reinforcement learning problems.
	How tested	CBL – group project
	Level	Create and apply
	Weight	33.3333…%
Goal 3	Goal	prove performance guarantees for the constructed solution strategies for reinforcement learning problems
	How tested	CBL – group project
	Level	Create, evaluate, analyze and apply
	Weight	33.3333...%

Figure 3.2: Assessment Matrix

research projects. The grade received by a student is fully determined by their research project.

The goals of the research project are the following :

1. Identify a research gap in the literature and pose a research question.
2. Develop a solution approach and implement their solution.
3. Prove mathematical guarantees for the solution.
4. Write an extensive report on their work.

Below I reproduce the Learning Objectives stated in Section 1.5

1. You will be able to formulate reinforcement learning problems mathematically.
2. You will be able to construct solution strategies for reinforcement learning problems.
3. You will be able to prove mathematical performance guarantees for the constructed solution strategies.

One can see that the first goal of the research project corresponds to LO1, the second goal of the research project corresponds to LO2 and the third goal of the research project corresponds to LO3. LO1 will be assessed based on how well the students formulate their chosen research problem. LO2 will be assessed based on the suitability of the implemented solution strategy. LO3 will be assessed based on the correctness of the mathematical analysis. Novelty in all three – the chosen problem, the solution strategy and the mathematical analysis, will contribute toward the assessment of the respective LO. The fourth goal of the research project corresponds to academic writing. Though academic writing is not an explicit learning objective for the considered course, it is a general skill a master's student is expected to possess.

All three goals have equal weight. These goals cover the levels of understanding, applying, analyzing, evaluating and creating from Bloom's Taxonomy (see Figure 3.1). The assessment matrix is given in Figure 3.2.

- **Formative Assessment**
  In the research project phase (i.e., weeks 4-8), each group is entitled to a weekly meeting with the lecturer. The discussions during these weekly meetings contribute to the formative assessment of students. In these discussions, the students consulted me on their research projects. The first couple of meetings were focused on determining the topic of the research project for each group. In the subsequent meeting. we discussed the issues the students were facing in their research projects and guided them toward the solutions to their respective issues. I varied the level of guidance according to the competence shown by the students – well-performing groups received hints so as to motivate them to arrive at the solution independently, while others received more detailed direction.

The help received during discussions helped the students to complete the research project in time. Six out of the total seven groups managed to submit their research project before the due date.

3.2 Meeting the Quality Criteria

I make use of *Ouriginal*, a plagiarism-checking tool that is available with TU/e credentials. Please see the charts given below where I illustrate how the assessment meets the criteria of validity, reliability and transparency.

Validity **The extent to which the test measures the envisaged learning objectives**	**Yes/No**	**Explanation**
The test is based on the test matrix.	No	It's a CBL project
The number and level of the test assignments correspond to the test matrix.	NA	
The level of proficiency of the test assignments relates to the learning objectives.	Yes	
The content of the test assignments relates to the learning objectives.	NA	There are no test assignments.
The assessment criteria are tailored to the assignments in terms of level and content.	Yes	The CBL project is designed according to the triangle of constructive alignment.
The assessment criteria assess (represent) the learning objectives.	Yes	(See above)
The number of assessment criteria corresponds to the importance of the learning objective (in accordance with the test matrix).		
The assessment form establishes a link between the learning objectives and the assessment criteria.	Yes	
The test reflects an assignment that may actually occur in professional practice.	Depends	This is a theoretical course so accordingly the test reflects an assignment that may occur in professional practice in academia or industry research in the theoretical foundations.

Figure 3.3: Validity of Assessment

Reliability **The extent to which the test leads to the same final conclusion under identical conditions.**	**Yes/No**	**Explanation**
Assignments		
The assignments allow the student to demonstrate his/her competences.	Yes	The students have lots of opportunities to identify a relevant and significant challenge and solve it in an independent manner.
The assignments differentiate between those students who perform well and those who perform less well (specificity).	Yes	The quality and quantity of results differentiate between students.
In the case of a group assignment, it is clear what the share of each individual student is.	Yes	We allow the students from each group to self-report the share of every student in the group. This, in addition to the lecturer's reflection, affects the grade.
The time that a student needs to complete the assignments corresponds to the time available.	Yes	All five weeks are available to students.
The assessment form includes the assessment criteria, assessment scale, marking system and cut-off score as well as the instructions for the assessor.	Yes	The rubric is made available to the students at the start of the course.
The assessment criteria are clear and specific and refer to the qualities of the achievement.	Yes	
The assessment criteria indicate the standard that the achievement must meet.	Yes	I encourage the students to seek clarifications.
The assessment criteria have been drawn up on the basis of consensus/calibration.	Yes	We revise it, if needed, according to student feedback and lecturers' reflections.
A clear assessment scale is used to evaluate the student's performance.	Yes	
There are enough assessment criteria to be able to comment on a student's performance.	Yes	
The instructions for assessors contain guidelines on making the assessment (scoring, marking, evaluating/grading).	Yes	
The assessors are qualified to administer the test and assess the student's performance.	Yes	
Several assessors are involved in assessing the test.	Yes	We discuss the grading with the lecturers involved in the course

Figure 3.4: Reliability of Assessment

Transparency **The extent to which the information about the test is clear to all involved.**	**Yes/No**	**Explanation**
The student is aware of the test assignment, assessment criteria, cut-off score and marking system, procedure (admissibility requirements) and conditions (place and time of the test, aids allowed and degree of independence) beforehand.	Yes	The rubric is made available to the students at the start of the course.
The test assignment is clear and specifies exactly what is expected of the student.	Yes	
The assignment sets out the conditions (place, time, aids, degree of independence) under which the assignment is carried out.	Yes	
The assessment criteria are transparent and make clear what is expected of the student (specific behavior) and what requirements are placed on the result.	Yes	
It is clear how the points are distributed (points scored), how the test is marked and what the cut-off score is.	Yes	
The assessment model includes room for feedback.	Yes	We seek feedback throughout the course and at the end.
The test includes formative assessments in preparation for the test.	Yes	We have weekly meetings during group projects during which we provide feedback on in-progress work

Figure 3.5: Transparency of Assessment

3.3 Assessment Forms

See Appendix I for an assessment form that we used for this course. The grading criteria are given on the last page of the assessment form. The grading criteria mimic the well-established criteria used to grade DS&AI MSc theses and hence I am convinced of their validity.

To depict how the assessment form can distinguish between groups, see Appendix J for an assessment form of a group that performed very well and K for an assessment form of a group that performed poorly. Following feedback from my UTQ coach Ms. Hester Morssink, I noticed that the assessment form does not explicitly mention novelty. Whereas, as noted in Section 3.1 novelty in all three – the chosen problem, the solution strategy and the mathematical analysis contributes toward the assessment of the respective LO. It should be noted that novelty is still being implicitly assessed without it being explicitly mentioned in the assessment form. To remedy this discrepancy, I will rename the "Results" section of the assessment form to "Problem, Solution and Results" and add

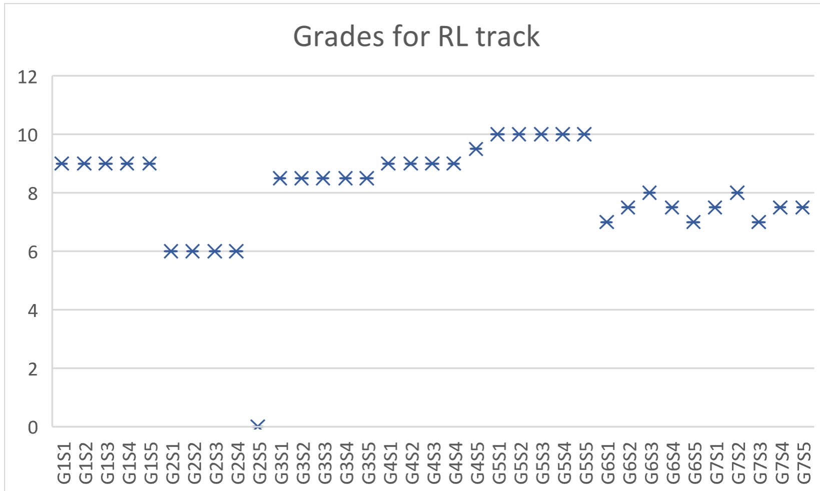


Figure 3.6: Grade for RL track – GiSj represents the j^{th} student in Group i. For G2S5, for ease of exposition, the grade is shown as 0 in the above chart even though their actual grade was "Incomplete"

novelty as a criterion in this renamed section.

3.4 Assessors

For the track of RL in the course 2AMM20, I am solely responsible for the assessment. However, this course consists of two more independent tracks namely – Exceptional Model Mining & Missing Data (EMM/MD) and Anomaly Detection (AD). For EMM/MD track, Dr. Wouter Duivesteijn and Rianne Schouten were responsible for teaching and assessment. For the AD track, Dr. Stiven Schwanz Dias was responsible for teaching and assessment. To make sure that all students are assessed in the same way across tracks, we hold a joint meeting after we have individually graded all the tracks (and before the grades are released). In this meeting, we discuss each group and see if all four of us are in unison with the assessment. In the last iteration of the course (2022-23 GS1), I changed the grade given to one of the groups from 7.5 to 8 after this meeting.

3.5 Results of Assessment

The grades obtained by the students can be seen in the chart given in Fig 3.6. In the chart, GiSj represents the j^{th} student in Group i. Only one student, G2S5, received an "Incomplete" grade as they did not participate in the course beyond the first few weeks. Apart from that, all the students received a grade higher than 6, the minimum required for a passing grade.

3.6 Analysis of Results

The histogram of grades achieved by the students can be seen in Fig 3.7. Quantitatively, I think the students performed as expected. 60%(21/35) of the students received moderate grades (between 7.5 and 9), while 20% of the students received low grade (less than 7.5), and nearly 18% of the students received high grade (more than 9). Qualitatively, I think there is room for improvement in argumentation skills and mathematical writing skills.

A notable deduction from the results is that all the students who completed the course received at least a passing grade. One of the students did not manage to complete to course and they received an Incomplete grade as a consequence. Following the discussions in the UTQ module on assessment and comments from Ms. Hester Morssink, I understand that the 100% pass rate might be contentious. Below, I provide probable reasons for the high pass rate for this course.

- Group project : Since the summative assessment was exclusively based on a group project, students could combine their individual skills and abilities in order to produce significant work. As each group consisted of five students, a straightforward calculation shows that the total amount of expected person-hours invested in one research project is 425 hours, i.e.: the full-time equivalent of a single person working for a quarter of a year (full year of work at TU/e is 1680 hours). Therefore the students had substantial time available to them to complete their project well.

- Continued Guidance : As detailed in Section 3.1, I provided guidance to each group during our weekly meetings. Via these weekly meetings, the students were able to identify and correct potential pitfalls in their projects early on.

3.7 Student Feedback on Assessment

In the last iteration of the course (2022-23 GS1), students were satisfied with this mode of assessment and they thought it is suitable for the course. See for example items 3,5,7,8,10,11 in the section titled "Final questions → What did you like about this course/project?" in the official students' evaluation enclosed in Appendix A.

3.8 Reflections on Assessment

For assessing master's students' theses, I make use of the standard departmental rubrics. I make these rubrics available to them in advance. As per the informal conversations I have had with the students, they found these rubrics to be suitable.

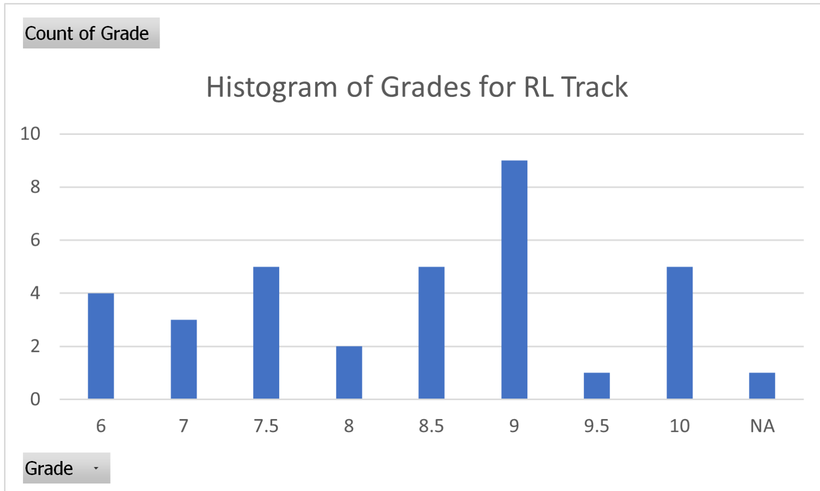


Figure 3.7: Grades for RL track – GiSj represents the j^{th} student in Group i. For G2S5, the grade is shown as 0 in the above chart for ease of exposition, even though their actual grade was "Incomplete"

As for summative assessment, I would impress upon the students to start the project as early as possible. This is following the feedback I received during informal chats with the students as well as some feedback given in the official students' evaluation enclosed in Appendix A. See for example items 3,6,9,14 in the section titled "Final questions → What would you like to improve in this course/project?" in the official students' evaluation enclosed in Appendix A.

I also noticed that the groups who spent considerable efforts in identifying the topics of their research project performed well. In particular, the group that scored the highest grade sent me a detailed list of potential topics before the first weekly meeting. In the first weekly meeting, we discussed each of these topics including their significance, difficulty level and suitability. Following our first meeting, the students shortlisted a couple of topics and performed a literature survey on those topics. In the second weekly meeting, we discussed those topics thoroughly and selected a topic for their group project. In the next iterations of the course, I would advise all the groups to follow the above.

In the last iteration of the course (2022-23 GS1), the most significant issue I faced with this assessment was how to assign grades to individual group members based on a collective grade for the group project. A solution we adopted was to seek a peer review from the students in which they are asked to rate the contribution of other students in their own group. This can be a single declaration on behalf of the entire group, in which each member writes an individual paragraph on the contribution of themselves and the group members. To allow for confidentiality, students may also (additionally) elect to individually write a peer review paragraph and submit it independently of their group. Based on these peer reviews, I adjusted the collective grade for the group project to individual group members, if needed. For example, see Appendix **??** for the relevant page from the assessment of one of the groups. The overall grade for the group project was 7.5. However, the peer reviews indicated that S2 had worked more than the others while S3 worked less than others. So S2 received 7.5+0.5, while S3 received 7.5-0.5, and the other three students received 7.5.

The reliance on peer reviews presented a problem when students were not open enough to acknowledge that one (or more) of the group members did not contribute equally. This could be because the students might not want to appear confrontational by deeming someone´s contribution as inferior while submitting a single joint declaration as a peer review. For example, in one of the groups, students said that all the members contributed equally. Although via my interaction with the group during our weekly meetings, one of the students was not participating in the discussion at all which led me to believe that they were not contributing equally to the project. However, I had to grade them all equally as it was not clarified beforehand that participation in the weekly meetings might affect the course grade.

To solve this problem, I propose the following changes.

1. Students will have to necessarily submit their peer reviews individually and their peer reviews will not be revealed to other students.

2. It will be made clear to them that participation in the weekly meetings might affect my assessment of their contribution to the group project (and thereby their individual grade).

Supporting Products

- Official students' evaluation enclosed in Appendix A.
- Assessment forms in Appendix I, J, K and L.

Chapter 4

Evaluating Teaching

4.1 Purpose and Methods of Evaluation

The purpose of the course evaluation was to appraise the quality of the course design, teaching and assessment. See the Evaluation Plan in Figure 4.1.

I used the following methods of evaluation.

1. Official student evaluation for the course 2AMM20 in Appendix A
2. Results from student questionnaire from the RL track in the course 2AMM20 in Appendix B.
3. Informal conversations with students.

The official student evaluation contains the standard questions set by TU/e. In order to receive feedback about issues not covered in the official student evaluation, I designed another questionnaire myself. The questionnaire was shared with the instructor of the UTQ course on evaluation. I made some changes to the questionnaire following their feedback. For example, I mentioned the expected time to complete the questionnaire in the beginning. Furthermore, I explained the answer choices in more detail and I provided an optional text-box to further clarify their answers.

4.2 Analysis of Evaluation and Conclusions

In this section, I will analyze the evaluation on the following three fronts – course design, teaching and assessment. The inferences I mention below are drawn from the official students' evaluation enclosed in Appendix A, the results from the student questionnaire enclosed in Appendix B and informal conversations with students.

What do you want to evaluate? *For example, one or more lectures or an entire course*	The entire course
What is the aim of the evaluation? *Do you want to improve your education based on the input and/or do you need the evaluation for example for a self-evaluation?*	1. To reflect upon my teaching 2. To improve my teaching skills
What do you need to know to reach the aim of your evaluation? *Do you want an impression of course quality? Or do you want feedback on a specific aspect (book, Canvas page, learning activity)?*	Course assessment is going to be through CBL according to the departmental vision. So, in light of that, I would like to know what changes could I make either in the content or in my teaching style to best achieve student expectations as well as expectations from the department.
How will you evaluate? *Course survey, lesson observation, discussion with students*	1. Course Survey 2. Student questionnaire 3. Peer evaluation
Who are the respondents? *Students, peers, QA officer*	Students and peers
Who is responsible for what? *Defining evaluation criteria / Distribution /Data analysis / Reporting results / Making changes / Monitoring changes*	QA officer responsible for defining evaluation criteria, distribution, and reporting results. I am partly responsible for evaluation criteria and I am fully responsible for making and monitoring changes.
When do you want to evaluate? *Before, during or after course/lesson*	Before the next iteration of the course
How will you present and distribute the results? *Present main results to colleagues, share all results. How do you communicate your evaluation results to your students?*	Students already have access to the official course evaluations (apart from open-ended questions), I believe. Apart from that, as Juul suggested, I will share the course evaluations from the previous iteration and also highlight the changes I made in the course due to that evaluation.
What do you want to do with the evaluation results? *Who will decide what to improve? And who will make the improvement(s)?*	I will decide what to improve (perhaps after discussing the evaluations with other faculty members) and I will make the improvements.

Sources:

- Making a course evaluation plan in 11 steps. Van den Berg and Gommer, 2017.
- Best Practices and Sample Questions for Course Evaluation Surveys, University of Wisconsin-Madison.

Figure 4.1: Evaluation Plan

4.2.1 About Course Design

- Level of difficulty: Most of the students thought the level of difficulty of the course was appropriate for a master's course. There was a concern raised by a student about the level of mathematical knowledge required to follow this course. I believe this is handled by the prerequisites set for the course which included elementary statistics and probability theory and comfort with applying mathematical tools.

- Educational setup and organization: Students liked that they worked on a research project which allowed them to perform novel research and write an academic report on it. Students also liked that they went through the whole research process from selecting a problem to writing a report. They also appreciated the freedom to decide the direction of their research project. Students believed that this whole experience would help them toward writing a research paper and indeed their master's thesis.

  Students also appreciated the weekly feedback sessions while they were working on their research projects.

  Some students said that workload distribution was imbalanced – light during the lectures phase and heavy during the research project phase. To address this issue, I will incorporate more student interaction during the lectures phase (See Section 2.3). I will also amend the workload distribution as outlined below in Section 4.3.2.

- Course materials: Students thought the course materials including the slides and the additional resources listed on Canvas were suitable and helped them to study for the course. In particular, students liked how the information was organized on slides with the help of a color pattern.

4.2.2 About Teaching

Feedback received from the students, Dr. Wouter Duivesteijn (see Appendix C) and Ms. Hester Morssink (see Appendix D) indicate that I am able to generate interest amongst students and I use a clear structure during teaching. Using the feedback given by Dr. Wouter Duivesteijn, I will also include a visual indicator of lecture progress on slides. I am also able to stimulate interaction however as Ms. Hester Morssink suggested in her feedback, I would repeat the questions asked by students in the class for the benefit of students attending the lecture online. Both Dr. Wouter Duivesteijn and Ms. Hester Morssink also mentioned that I tend to look at the screen a bit too much. I will work to remedy this concern.

4.2.3 About Assessment

Students agreed that producing novel research and writing an academic report on it was the suitable assessment for the course.

4.2.4 Learning Environment

Students thought that I was able to create an inclusive and accessible learning environment during the course.

4.3 Reflections and Improvements

4.3.1 What worked well

- The course setup and method of assessment.
- Regular feedback during the research project Phase: In the official students' evaluation and informal conversations with students, they indicated that they liked the weekly feedback sessions during the research project Phase. Such regular feedback helped them to develop their research project in a well-informed manner.
- Freedom to choose the topic of the research project.
- Effectiveness of lectures and slides: Students appreciated that the lectures and the accompanying slides helped them to understand the subject well. They liked the way all the information was organized on the slides (colors, variables mathematical formulations, and equations) and the use of real-life examples.

4.3.2 Points of Improvement

- More Interaction with students: I am going to make a number of changes to the course to provide more interaction opportunities for the students. See Section 2.3 for concrete improvements.
- Amending the course timeline to encourage the students to start their research project earlier: Some students noted that the workload during the course was imbalanced – in the initial few weeks of the course, the workload was light, and in the last five weeks (i.e., the research project phase), the workload was quite heavy. So in the next iteration of the course, I would encourage students to start the research project as early as possible.
- Managing group dynamics during the research project phase: During informal conversations with the students, a few of them mentioned that they were not able to form an effective and balanced working relationship among the group members. In the next iteration of the course, I would like to identify such problems in group dynamics and suggest a viable solution to the students.

Supporting Products

- Official student evaluation for the course 2AMM20 in Appendix A
- Results from student questionnaire from the RL track in the course 2AMM20 in Appendix B.
- Feedback on Teaching Activity from Dr. Wouter Duivesteijn in Appendix C.
- Feedback on Teaching Activity from Ms. Hester Morssink in Appendix D
- Feedback on Teaching by a Student in Appendix E

Chapter 5

Professionalization

In this chapter, firstly I will explain my teaching vision in Section 5.1. Next, I will comment on the various University bodies relating to my teaching in Section 5.2. In Section 5.3, I will articulate my collaboration with colleagues and their feedback. I will conclude the chapter with reflections on my role and function as a teacher and plans for improvement. As supporting products for this chapter, I have enclosed the self-assessment form at my intake meeting and the current self-assessment.

5.1 Teaching Vision

- Facilitating learning catered to students' background and their academic level: While teaching a course, I try to understand students' background and motivations in order to adapt my teaching content and teaching style. For example, if I am teaching machine learning to second-year Computer Science master's students, I would incorporate strategies from the challenge-based learning paradigm via first teaching them the Big Idea, then assisting them to investigate and formulate a challenge and finally guiding them toward scientific solutions to the identified challenge. On the other hand, if I am teaching machine learning to second-year Industrial Design bachelor's students, I would use more conventional instructional teaching with regular assignments to develop their basic understanding of the topic.

- Inclusive and accessible education: Research has shown that students perform better when diverse socio-cultural backgrounds of students are recognized and embraced within educational institutions. It is also worth noting that an inclusive approach not only improves the performance of students from underrepresented communities, but that of other students as well [10]. In order to provide diverse and inclusive teaching and supervision, I act according to the guidelines shown in Table 5.1

Principle	Activities
Positive class climate	Learning names, in-class surveys and activities
Explicit expectations	Clear assessment criteria, timely feedback
Diverse course content	Use of multiple and diverse examples
Accessible course	Use of dyslexia-friendly fonts (e.g. arial)
Commitment to inclusion	Self-inventory of biases, ways to overcome them

Table 5.1: Guidelines for inclusive and accessible education

5.2 University Bodies Relating to My Teaching

Firstly, as mentioned in Chapter 1, we took the departmental consideration the departmental vision – the end goal of this course is that students will make a research contribution : they will go beyond the current state of the art.

As part of the UTQ course on evaluation, I also had an extensive meeting with the Quality Assurance officer for the Department of Mathematics and Computer Science, Ms. Nataly Alarcon Cepeda. In this meeting, she explained the procedure for creating the official evaluation form for a course. We discussed various avenues in which the concerned lecturer can provide their input and modify the evaluation form to their requirements.

During my conversations with other lecturers, I also learned the operations of the Examination Committee. In particular, a lecturer spoke about the scenario in which a student who had failed a CBL course contested the decision in front of the Examination Committee. In this case, detailed written notes about the student's performance throughout the course helped the lecturer justify their decision. As my course was also based on the CBL paradigm, I too kept detailed notes about every group's performance throughout the research project phase.

5.3 Collaboration with Colleagues and Their Feedback

I have been a part of the teaching team for the following courses at TU/e.

- JBI030 Data Mining
- 2AMM20 Research Topics in Data Mining
- DBM140 Embodying intelligent behavior in social context
- 2AMC15 Data Intelligence Challenge

As a part of these teaching teams, I have collaborated with faculty members from the Department of Mathematics and Computer Science and the Department of Industrial Design. See the letters from Dr. Wouter Duivesteijn in Appendix M and Dr. Emilia Barakova in Appendix N. Dr. Emilia Barakova was the responsible lecturer for course DBM140 which I co-taught in Q1 2022-2023. I have taken their feedback on board and I have identified plans for improvement (e.g., enrolling in TEACH650-15 Powerful Communication and Voice Techniques as mentioned below).

5.4 Reflections on My Work and Function as a Teacher and Plans for Improvement

I think my skills across all of the five competences have improved since I started the UTQ trajectory. See the two versions of the self-assessment form in Appendix O and P. The self-assessment form in Appendix O was completed on October 27, 2022 and the self-assessment form in Appendix P was completed on September 4, 2023. Below I expand on the improvements and further plans relating to all five competences.

- Designing teaching : I think I have improved a lot in designing teaching based on the principles of constructive alignment. I believe this is what I was doing instinctively even before learning about the paradigm of constructive alignment. However, learning this paradigm, discussing it thoroughly with the UTQ instructors and other lectures, and then practicing it has given me a structured way to design teaching. This helps me to design teaching in order to achieve the set learning objectives and test if they have been achieved. I have also been able to understand the logistics of teaching online and in-person classes and appreciate the differences between them. My experiences have given me insight into the feasibility and suitability of certain teaching methods in various scenarios. As for designing a course that fits the overall curriculum, I think I could do better with more teaching experience.

- Teaching and supervision : Preparing an educational meeting and conducting it are the activities that I enjoy the most. I believe I have improved the most on these two fronts. I still experience some stage fright before a large audience, especially before the first class of a course. However, I have worked on it for the past many years and I will continue to do so. I sometimes struggle with the degree of clarity and distinctness of pronunciation in speech. To deal with these issues, I have enrolled in the following course – TEACH650-15 Powerful Communication and Voice Techniques.

- Assessment : In terms of assessment, I now understand better how to design and implement the assessment for specific learning objectives. The course assessment of the course that I was teaching consisted of only seven groups. Therefore, I was unable to practice the statistical analysis techniques that I learned in the UTQ course on assessment. However, I believe that they will be useful when I teach a larger course.

- Evaluating teaching: I believe I have improved fairly in this activity. I would continue to evaluate my own as well as others' teaching in order to practice the analysis techniques.

- Professionalization: I have a much better understanding of what it means to be a professional teacher, the potential pitfalls and improvement techniques. I understand the theory behind various teaching competences via the UTQ courses and I have been able to apply it in practice in my teaching.

There are some other salient issues that I would like to mention here:

- Dealing with conflicts: In my course, one of the groups complained to me that a particular member was not contributing to the group project. By conducting a reconciliatory meeting with that group, we were able to arrive at a possible solution acceptable to all the involved parties.

- Balancing different professional roles: This is something I would like to improve on. In the quartiles that I teach, I struggle to find enough time to dedicate to research.

- Clear communication and structure: I am at my best when the communication with the stakeholders is clear and when the structure for the collaboration is clearly identified and communicated early on. This was the case with one of the courses that I was involved in. As a consequence, I enjoyed teaching the course and I was able to perform at a high level. On the other hand, clear communication was a challenge in another course that I was involved in. As I was not the responsible lecturer, I was unable to take the lead and resolve these issues. If I am in this situation again, I will be more persistent in establishing a working collaboration that suits everyone in the team.

Supporting Products

- Official students' evaluation enclosed in Appendix A.
- Support letter from Dr. Wouter Duivesteijn in Appendix M.
- Support letter from Dr. Emilia Barakova in Appendix N.
- Self-assessment form completed on October 27, 2022 in Appendix O.
- Self-assessment form completed on September 04, 2023 in Appendix P.

Bibliography

[1] Richard E. Mayer. *Multimedia Learning*. Cambridge University Press, 2001.

[2] Paul Main. Think, pair, share: a teachers guide. *Structural Learning*. URL `https://www.structural-learning.com/post/think-pair-share-a-teachers-guide`.

[3] Lukas Mundelsee and Susanne Jurkowski. Think and pair before share: Effects of collaboration on students' in-class participation. *Learning and Individual Differences*, 88:102015, 2021. ISSN 1041-6080. doi: https://doi.org/10.1016/j.lindif.2021.102015. URL `https://www.sciencedirect.com/science/article/pii/S1041608021000522`.

[4] Patrick Henry Winston. How to speak. University Course, 2018. URL `https://ocw.mit.edu/courses/res-tll-005-how-to-speak-january-iap-2018/`.

[5] Theo Wubbels, Mieke Brekelmans, Perry Brok, and Jan van Tartwijk. An interpersonal perspective on classroom management in secondary classrooms in the netherlands. 2006. URL `https://api.semanticscholar.org/CorpusID:152293363`.

[6] Jiong Li and Pratik Gajane. Curiosity-driven exploration in sparse-reward multi-agent reinforcement learning. In *the 16th European Workshop on Reinforcement Learning (EWRL)*, 2023.

[7] Ronald C. van den Broek, Rik Litjens, Tobias Sagis, Luc Siecker, Nina Verbeeke, and Pratik Gajane. Multi-armed bandits with generalized temporally-partitioned rewards. In *the 16th European Workshop on Reinforcement Learning (EWRL)*, 2023.

[8] Danil Provodin, Pratik Gajane, Mykola Pechenizkiy, and Maurits Kaptein. The impact of batch learning in stochastic linear bandits. In *2022 IEEE International Conference on Data Mining (ICDM)*, pages 1149–1154, 2022. doi: 10.1109/ICDM54844.2022.00146.

[9] Rosa van Tuijn, Tianqin Lu, Emma Driesse, Koen Franken, Pratik Gajane, and Emilia Barakova. Weheart: A personalized recommendation device

for physical activity encouragement and preventing "cold start" in cardiac rehabilitation. In José Abdelnour Nocera, Marta Kristín Lárusdóttir, Helen Petrie, Antonio Piccinno, and Marco Winckler, editors, *Human-Computer Interaction – INTERACT 2023*, pages 191–201, Cham, 2023. Springer Nature Switzerland. ISBN 978-3-031-42286-7.

[10] Laura Celeste, Gülseli Baysu, Karen Phalet, Loes Meeussen, and Judit Kende. Can school diversity policies reduce belonging and achievement gaps between minority and majority youth? multiculturalism, colorblindness, and assimilationism assessed. *Personality and Social Psychology Bulletin*, 45(11):1603–1618, 2019. doi: 10.1177/0146167219838577. URL `https://doi.org/10.1177/0146167219838577`. PMID: 31014198.

Appendix A

Official Students' Evaluation

Research Topics in Data Mining (2AMM20) 2022/2023 Q1

Report composed on 28-11-2022

Course name (2AMM20_2022_1)	Research Topics in Data Mining (2AMM20) 2022/2023 Q1
Evaluation name	Research Topics in Data Mining (2AMM20) 2022/2023 Q1
Labels	Mathematics and Computer Science, Q1, 5ECTS, GS
Evaluation start- and end date	13-11-2022 t/m 27-11-2022
Amount of respondents	32 from a total of 110 (29%)

Table of contents

Average scores

Below are the average scores. These averages are composed of all results on all questions, with the exception of the questions with the scales "Yes / No" and "Open question", and questions in which the set of questions states that they may not be included in the average.

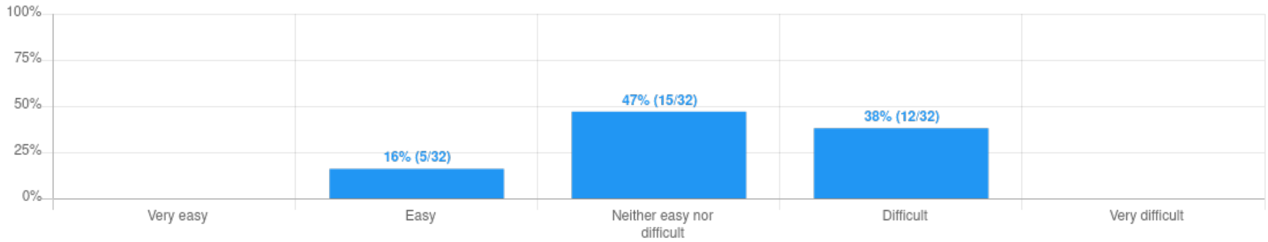
Average score total	Average score course	Average score teacher	Your average score
7.9	7.7	9	8.9

Questions

General

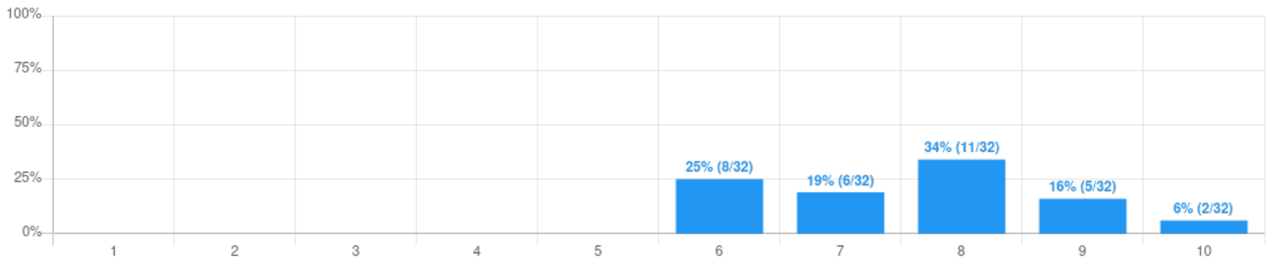
Overall, how would you describe the level of difficulty in this course/project? — Average score: 3.2

Scale: Very easy to very difficult | σ 0.7 | Number of given answers: 32 — Neither easy nor difficult



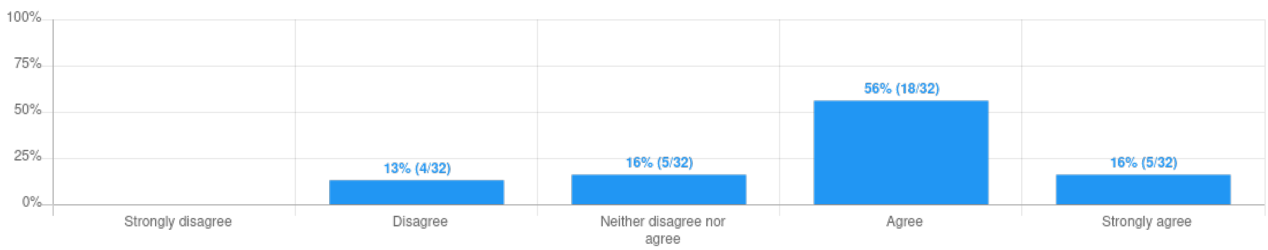
On a scale of 1 to 10, how would you rate this course/project? — Average score: 7.6

Scale: 1 to 10 (points) | σ 1.2 | Number of given answers: 32



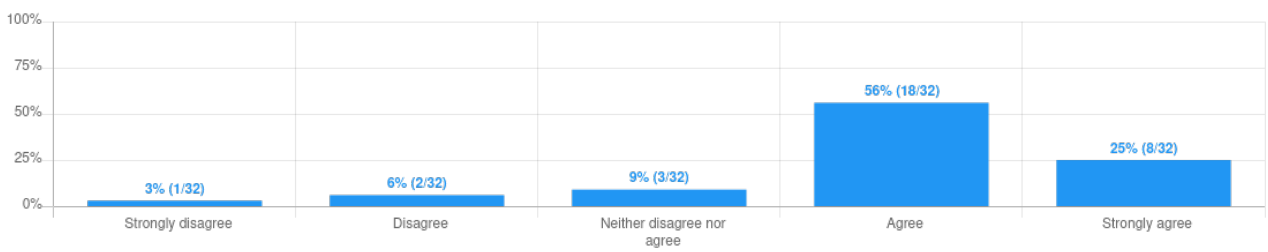
The educational setup (e.g. structure, content, teaching/learning methods, level, and coherence) worked well and was suitable for this course/project. — Average score: 3.8

Scale: Disagree to agree | σ 0.9 | Number of given answers: 32 — Agree



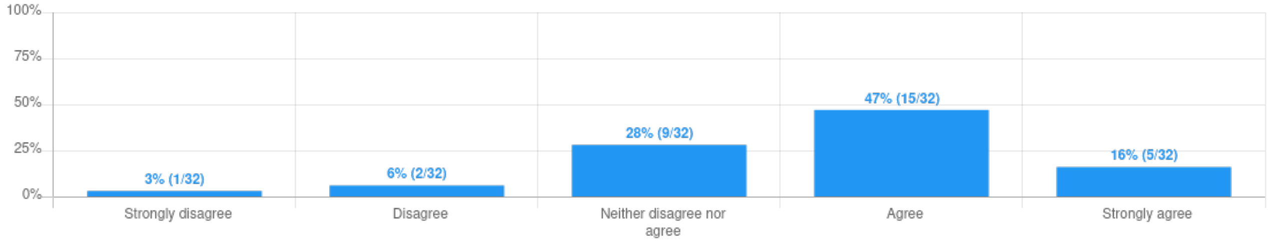
The course/project was well organized. — Average score: 3.9

Scale: Disagree to agree | σ 0.9 | Number of given answers: 32 — Agree



The course material was clear and motivated me to study for this course/project. — Average score: 3.7

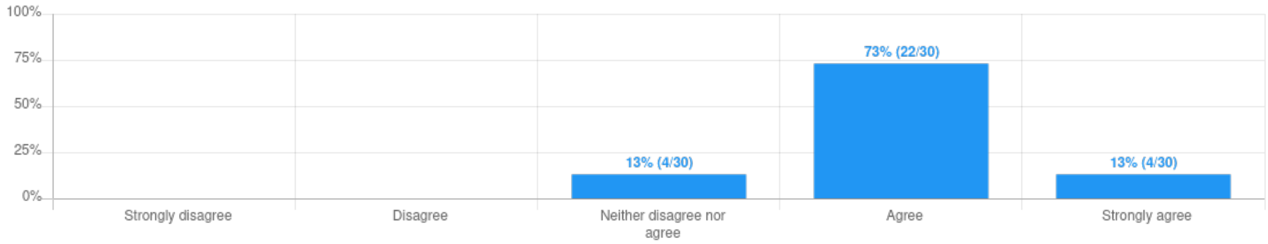
Scale: Disagree to agree | σ 0.9 | Number of given answers: 32 — Agree



The assessment of this course/project was appropriate (e.g., methods used, relevance and clarity of the questions/assignments).

Scale: Disagree to agree | σ 0.5 | Number of given answers: 32 | Total Don't know / Don't want to answer: 2

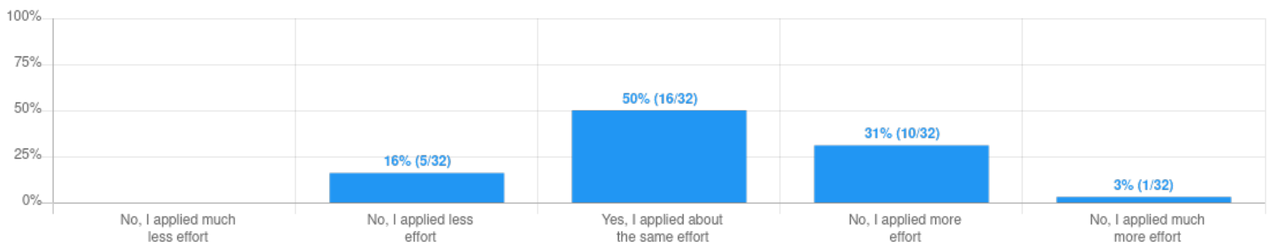
Average score: 4
Agree



Did the effort you applied to complete this course/project correspond with the number of credits? (1 ECTS = 28 hours; 5 ECTS = 140 hours)

Scale: Much less to much more effort | σ 0.7 | Number of given answers: 32

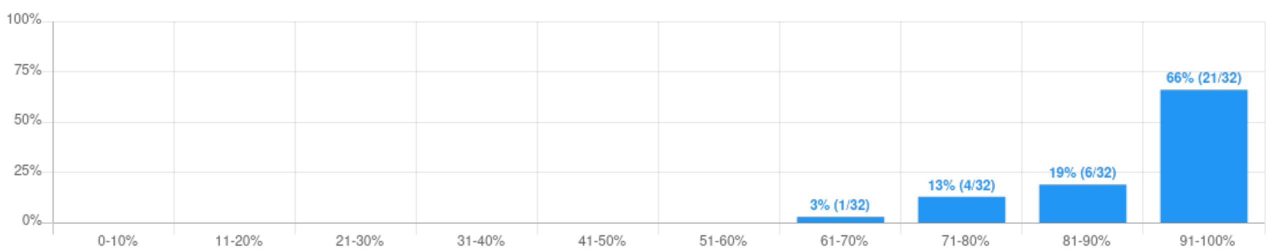
Average score: 3.2
Yes, I applied about the same effort



What percentage of the teaching sessions did you attend?

Scale: Percentage ten steps | σ 0.8 | Number of given answers: 32

Average score: 9.5
91-100%



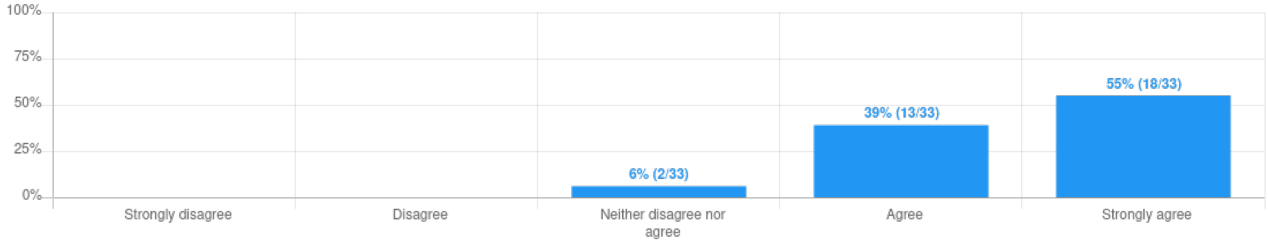
Teacher questions

These answers are about all the teachers

The lecturer explained the content in a clear and comprehensive way.

Scale: Disagree to agree | σ 0.6 | Number of given answers: 44 | Total Don't know / Don't want to answer: 11

Average score: 4.5
Strongly agree



Final questions

What did you like about this course/project?
Scale: Open question | 22 answers by 32 respondents

1. *The reinforcement learning track requires math knowledge that is too deep*
2. *This course has a very free structure. You can choose your own track and after that you can choose your own project. The content is also very interesting*
3. ** Weekly Feedback Sessions * That we went (more or less) through the whole research "cycle". First iterating on our research problem, ... * Freedom to decide the direction of our research project on our own*
4. *The lessons of the first weeks were clear and interesting (for reinforcement learning). Pratik was nice and clear when explaining.*
5. *The project allowed us to pratice doing actual research and writing an academic paper.*
6. *The lectures were pretty fun to attend without having pressure. And I learned a lot about how to write/read a paper. The frequent feedback that my group gets was useful.*
7. *Doing research in a novel area*
8. *I really liked the project structure and the set up of the course*
9. *Prepares me well for my master's thesis*
10. *This course prepared the students toward research approach and research paper writing in the field of data mining.*
11. *I liked that we can experience what its like to write a research paper*
12. *The lecture for "what is EMM" was perhaps the best (certainly top5) lectures I have ever attended. The "write a paper" setup was also the most reasonable one for such an academic theory course*
13. *As a course, it was very insightful for me as I had no previous experience with academic writing. Thus, I learned many new things, both in technical and a more theoretical way, that I am sure they will help me in the rest of the master's studies.*
14. *I enjoyed working on the research project. Even though the lectures in the first part were few, they contained everything we needed for the research project. There were also many papers provided for helping with the related work section.*
15. *That you have to write a paper in a team and work on new problems*
16. *the first part when we got lectures*
17. *I learned how academic conferences work and participated in a specific research topic. I think it's essential for a master's student to learn about them.*
18. *The structure of the group project was fun, by students supplying work to the lecturer who'd then give feedback as if we were working for some employer. This organization encouraged us to put a lot of effort into the course.*
19. *I like the topic of Learning. I enjoyed a lot the lectures. The topic seems difficult, but I think the teacher did an amazing job with the way all the information is organized on the slides (colors, variables, mathematical formulations, and equations), and all the real-life examples made the topic interesting. I think the teacher explained the content in a clear and comprehensive way and made it seem easier than it actually is. This course increased my interest in Reinforcement Learning. And if there is ever a course on this topic, I think Pratik would be great for it.*
20. *This course is set up in a way unlike any course I have taking before. The one lecture Wouter Duivestein gave in person about EMM was one of the best lectures i have followed. Maybe reduce the paper set phase. It was nice but more time for the report/project would be nice. The research question can make or break the project.*
21. *Lots of freedom to choose your own topics.*
22. *I liked that it was practical. Most courses aren't practical in TU/e and I think that is not good for those who want to go to the industry.*

What would you like to improve in this course/project?
Scale: Open question | 17 answers by 32 respondents

1. *Speaking about reinforcement learning it would have been nice to have some track to work on given by the teacher*

2. *How the final project is handled. I got put in a group with 3 other students that got in the course late (totally understandable) since we were all too late to find a group. During the second meeting, I notified our teacher that things were not working out within our group and that we wanted to split, but we were told that we should work it out ourselves and do it with the 4 of us. 2 of us were not able to work on the code since they have macbooks, but they refused to work on uni's computers, so everything had to be done on my laptop. For some reason I did not get backed up by the teacher when I said they should use the computers at uni, as it should not be my problem that they cannot work on it.*

3. *I feel the lecture period in the beginning could be 1 or 2 weeks less to have more time for the research paper.*

4. *It might be that I misunderstand the purpose of the course, but for the most part I feel as if I now know more about a very specific part of Multi-armed bandits, that I am unlikely to use again, instead of a more general understanding of reinforcement learning.*

5. *It feels like the first two projects were not very impactful and the time invested in those could have been used differently*

6. *Perhaps the project phase could start a little bit earlier, as the workload feels a bit outbalanced. The first weeks, the workload is very low compared to the workload in the latter weeks. As interesting as the lectures were, I don't directly see the use of them for our particular project.*

7. *The size of the groups should have been a lot smaller.*

8. *I do not have any comments. It was a nice course.*

9. *More time to work on the project, or maybe that the groups are created earlier*

10. *I think this course is good and well organized. Currently, I do not have any idea of improvement.*

11. *The time is very short and we did not have a clear idea on what scope we should take on for the paper*

12. *The workload distribution change between the early parts of the course vs. after the project phase has started was....quite overwhelming. I wish the distribution could have been a bit more smooth*

13. *I think a nice little way to improve this course is to include a lecture about the technical aspects of academic writing between the subject-matter lectures of each track and the beginning of the writing. This will help many students who have no prior experience of academic writing. It would be a nice opportunity to explain things such as the importance of citations, the language and the symbols one should use while writing an academic paper, the structure etc.*

14. *The time we had in order to write the paper was too little. I think it is better to already start with the paper while doing the lectures. And it will maybe help to set strict deadlines in the middle of the course to make sure that people dont fall behind.*

15. *The group work was very difficult to organise, and some group members took the opportunity to coast through the course*

16. *I would sort of make this course a prerequisit for entering the master's program. During the project, it became clear that some students had never read scientific literature before, or that some students had never done academic writing before. As I think these are 2 very important qualities when following a master's degree (or a bachelor's degree for that matter), I would argue that students were to be evaluated on the aspects that are covered in this course, before actually being able to enroll in the course/program.*

17. *Expand more topics like MLOps. Concept drift, etc. Things that are new to the industry and that I am sure the professors would be great in the area too.*

Appendix B

Results of Online Questionnaire

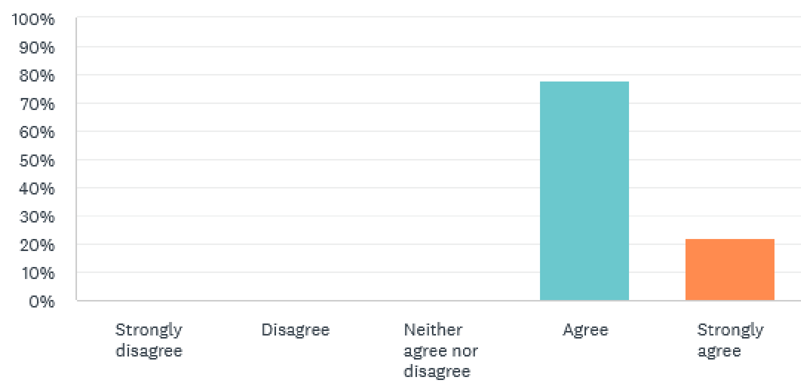


Figure B.1: Percentage of responding students agreeing with – "The course materials supported the content well."

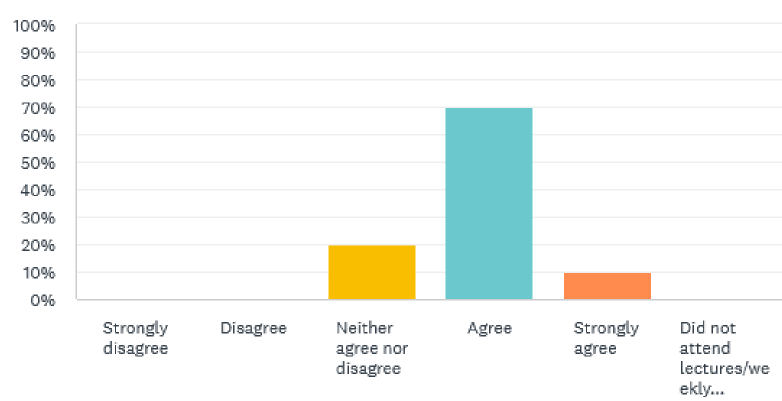


Figure B.2: Percentage of responding students agreeing with – "The lecturer fostered an inclusive and accessible learning environment."

Appendix C

Feedback on Teaching Activity from Dr. Wouter Duivesteijn

Feedback form on teaching activity

Participant: Pratik Gajane
Observer: Wouter Duivesteijn
Date: September 14, 2022 (lecture); April 04, 2023 (feedback)
Title of teaching activity: Lecture “Thompson Sampling for Bandits” in course 2AMM20

Observed teacher behaviour		
Introduction	**Observed?**	**Description of observed behaviour & feedback**
Generates interest	Yes	From 6:52 to 20:03, makes clear shift between the two main statistical paradigms (frequentist versus Bayesian approach). Clearly delineates what has come before (frequentist) and what is about to come (Bayesian), while encasing these concepts in the red through-line that threads the lectures together.
Describes learning outcomes	No	This lecture would not be the place to make this explicit. The learning objectives of 2AMM20 include identifying and resolving a research gap in scientific literature. This lecture serves to get students up to speed with the current state-of-the-art in reinforcement learning, which is necessary for achieving the learning objectives but cannot be directly tied to specific learning objectives.
Clarifies relationship between learning outcomes, teaching activities and assessment	No	Assessment is simply not the focus of this particular lecture.
Activates prior knowledge	Yes	Explicit call-back to lessons learned in previous lectures from 0:18 to 6:47.
Body		
Uses a clear structure	Kind of	Pratik has a clear structure planned out for the lecture, and sometimes the slides also indicate it, but there are times when there is no visual indicator of where in the story we are at the moment.

Uses different methods to retain student interest and attention	No	There is not a big variety of methods. I am not convinced that there needs to be such a variety.
Used methods and activities support students in achieving the learning objectives	Yes	Almost by definition: by understanding the state-of-the-art reinforcement learning research, students will be better able to identify and bridge a research gap.
Closing		
Summarizes the main topics of this lesson	No	I observed only the first half of a lecture, which didn't really do a summarization at the end. There's just a lookahead towards what we will do in the next half.
Refers back to the learning outcomes and/or assessment	No	I observed only the first half of a lecture, which didn't really do a summarization at the end. There's just a lookahead towards what we will do in the next half.
Reflects on the lesson	No	I observed only the first half of a lecture, which didn't really do a summarization at the end. There's just a lookahead towards what we will do in the next half.
Didactical skills		
Activates students	Yes	At 33:15, in response to a student question, Pratik pulls up the blackboard to explain some concept. This in itself is already activating: students will wake up to see what on earth is going on right now.
Stimulates interaction	Yes	Explicit request for student feedback at 20:57. "Is that clear?" at 23:15; checks whether the student understood the answer to the question the student asked.
Motivates students (to take charge of their own learning process)	Yes	See request mentioned above; the motivation pays off at 22:40 when the audience feels empowered to ask a question which Pratik then answers, followed by another question at 23:00.
Differentiates for differences between learning styles, cultures and/or functional impairment.	No	Couldn't see anything specific to this point in this lecture.
Uses ICT/blended learning to optimize learning	No	This lecture did not include specific ICT/blended learning. Instead, it's a traditional plenary lecture.

Incorporates effective feedback	No	Couldn’t see anything specific to this point in this lecture.
Teacher’s presence		
Has eye contact with students	Yes	Makes eye contact with people in the audience all the time. Inbetween, tends to look back at the screen a bit too much; would be better to look the smaller monitor that is also provided in the lectern at which you can choose to stand; that way, your head is still tilted towards the audience. It’s tricky though; standing behind the lectern runs the risk of being stilted, while the position Pratik takes enables a more active body language towards the audience. So even though there are drawbacks to the made choice, there are also clear benefits.
Easy to approach	Yes	Pratik actively works to make the lecturer-student distance smaller by asking audience questions and allowing the audience to ask questions at any time. Students make use of it too.
Shows enthusiasm	Yes	The entire lecture simply radiates that Pratik is enthusiastic about the topic, and enthusiastic about telling the students more on the topic. Students will pick up on that enthusiasm, and that stimulates learning.
Speaks clearly (f.e. articulation, speed, intonation)	Yes	In everyday office life, Pratik speaks a whole lot faster than he does in a lecture. This makes sense. In the lecture, you want to take more time to let your words sink in well in the students, and you want to let students get used to your accent (time and time again). The slower speed of speech will make the language used much more interpretable by the students, so this is great.
Makes good use of movement and gestures	Kind of	Tends to use his hands to stress certain narrative beats, which is good. Could make more use of similar body-language tricks.
Other comments & feedback:		

In closing, even if it's only the closing of the first half of a lecture, it may be worth it to explicitly summarize what the students have learned so far. Give them the highlights of the first half before you send them off into the coffee break, to ensure that they retain more of the material. In my lectures, I have sometimes put an explicit "What have we learned so far" slide in the middle, but this is a risky move, since it forces you to have the break right there. If you go a little faster or slower, this makes things awkward. Still, in the current form, the last few lines before the break feel rather anticlimactic.

Appendix D

Feedback on Teaching Activity from Ms. Hester Morssink

Feedback form on teaching activity

Participant: Pratik Gajane
Observer: Hester Morssink
Date: 27 – 3 – 2023 (video-observation)
Title of teaching activity:

Observed teacher behaviour		
Introduction	**Observed?**	**Description of observed behaviour & feedback**
Generates interest	Y	Beginning of lecture could be a bit more "lively"
Describes learning outcomes	N	Not mentioned explicitly in slides
Clarifies relationship between learning outcomes, teaching activities and assessment	N	Not explicitly in this lecture
Activates prior knowledge	Y	Does recap of previous lectures. Maybe it would be possible to let the students do the recap? To involve them more.
Body		
Uses a clear structure	Y	Structure is made clear in lesson plan, but not explicitly to students? In general structure is easy to follow.
Uses different methods to retain student interest and attention	N	Not explicit
Used methods and activities support students in achieving the learning objectives	N	LO are formulated as "recall, solve and explain". However, the lecture is a one-way explanation. Lecturer can use more interactive questions or assignments to check understanding and learning outcomes.
Closing		
Summarizes the main topics of this lesson	N	Only seen first half of the lecture. The lecturer pauses for a break, and no recap or check understanding is done.
Refers back to the learning outcomes and/or assessment	N	NA. End of lecture is not observed

Reflects on the lesson	N	NA. End of lecture is not observed.
Didactical skills		
Activates students	N	Not explicitly. There is room for questions. Point for feedback: please repeat question from student for the rest of the audience (also for audience at home?)
Stimulates interaction	N	Not explicitly. Lecturer gives some room for questions, but this could be done more. The recap could also be done with room for questions .
Motivates students (to take charge of their own learning process)	N	This is quite a teacher centered lecture. Difficult to see how students are motivated.
Differentiates for differences between learning styles, cultures and/or functional impairment.	N	Not observed in this lecture.
Uses ICT/blended learning to optimize learning	Y	Uses clear slides. Maybe other tools could be used for recap, interactive questions, check students understanding, etc
Incorporates effective feedback	N	NA
Teacher's presence		
Has eye contact with students	Y	Lecturer looks at the screen a lot. I think more contact with the audience could be possible
Easy to approach	Y	Showing from the fact that students ask questions during lecture.
Shows enthusiasm	Y	Lecturer explains well and clearly is enthusiastic about the topic
Speaks clearly (f.e. articulation, speed, intonation)	Y	
Makes good use of movement and gestures	Y	Could be done more maybe
Other comments & feedback:		

We discussed this lecture with Pratik, and he indicated he recorded this lecture before taking part in the UTQ modules. He indicated himself that he would do much more to activate students and create interaction.

Appendix E

Feedback on Teaching by a Student

From: [redacted]
To: Gajane, Pratik
Subject: Feedback on teaching and supervision
Date: Friday, December 2, 2022 10:07:07 AM

Dear Pratik,

The link for feedback has expired. However, I filled in the course evaluation of the course, and from what I can remember, I might have written my feedback specifically for the RL track in the open question at the end. Overall, I would like to thank you for teaching the RL part of the course and for your supervision. I also want to highlight that Reinforcement Learning seems like a very difficult topic for people with a weak background in mathematics, but I felt that you taught us everything step by step and made it easier for us to understand it.

Kind regards,
[redacted]

Appendix F

Clarifying Questions asked by Students

From: Gajane, Pratik
To: [redacted]
Subject: RE: Question about last part of proof Chernoff-Hoeffding bound [2AMM20]
Date: Thursday, September 15, 2022 10:49:00 AM
Attachments: image001.png
image002.png
image003.png
image004.png
RL_part-of-proof (PG comments).pdf

Hi [redacted],

You got all the steps right but the log is to the base e and not 10 and exp(x) is just another way of writing e^x. Please see my comments on your proof in the attachment.

p.s. It's great to see that you are actually verifying the math. Keep it up!

Kind regards,
Pratik

From: [redacted]
Sent: Wednesday, September 14, 2022 9:45 PM
To: Gajane, Pratik <p.gajane@tue.nl>
Subject: Question about last part of proof Chernoff-Hoeffding bound [2AMM20]

Dear Pratik,

I have a question about the last step in the proof of the Chernoff-Hoeffding bound. Slide 34 of lecture 2 states:

$$\epsilon = \sqrt{\frac{2\log(1/\delta)}{n}} \rightarrow \delta = \exp\left(-\frac{\epsilon^2 n}{2}\right)$$

I am not sure why we can say that:

$$\delta = \exp\left(-\frac{\epsilon^2 n}{2}\right)$$

If I start with:

$$\epsilon = \sqrt{\frac{2\log(1/\delta)}{n}}$$

And I try to rewrite it in the form δ = ..., then I get $10^{\frac{-\varepsilon^2 n}{2}} = \delta$

(See appendix for the steps)

Could you please explain it to me?

Kind regards,

[redacted]

Appendix G

Feedback on Supervision

G.1 Feedback 1

EINDHOVEN UNIVERSITY OF TECHNOLOGY

Department of Mathematics and Computer Science
De Groene Loper 5,
5612 AZ, Eindhoven
The Netherlands
P: +31 40 247 2733 (secretary)
e-mail: d.provodin@tue.nl

March 3, 2023

To whom it may concern,

I am pleased to provide feedback for my closest collaborator Dr. Pratik Gajane. I have actively collaborated with Pratik since winter 2021 when he joined the Eindhoven University of Technology as a Post Doc. Over the past two years, his expertise and commitment have been integral to my personal and academic growth.

I appreciate the time Pratik has taken to provide thoughtful feedback on my work and the opportunities he has given me to expand my knowledge in the field. His constructive criticism and encouragement have been vital in my growth as a researcher. Based on my personal impression, I believe that Pratik, with his academic background and strong leadership skills, would be an excellent professor.

Kind Regards,

Danil Provodin,
PhD candidate

G.2 Feedback 2

"Special thanks to my tutor Dr.Pratik Gajane, who is highly supportive and patient in communicating and guiding me since this project's initialization and development. Moreover, He encourages me to publish this project in the workshop and gave invaluable help for the publication in EWRL 2023."

Appendix H

Lecture Slides

Lecture 3 - Thompson Sampling for Bandits

Pratik Gajane

September 14, 2022

2AMM20 Research Topics in Data Mining
Eindhoven University of Technology

A Quick Recap of Lecture 1

- Introduction to reinforcement learning (RL).
- Mathematical formulation of a RL problem.
- Formulating RL with multi-armed bandits and its variants.
- Formulating RL with Markov decision processes.

Recap Lecture 2 : Stationary stochastic bandits

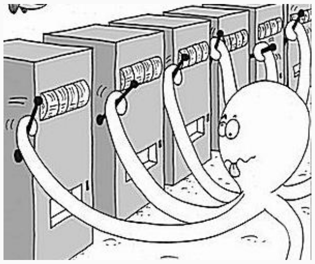


Image source : *Microsoft research*

- At each time step t, the agent selects an action $i(t)$ and then receives a numerical reward $r(t) \sim X_{i(t)}$ with mean $\mu_{i(t)}$.

Recap Lecture 2 : Stationary stochastic bandits

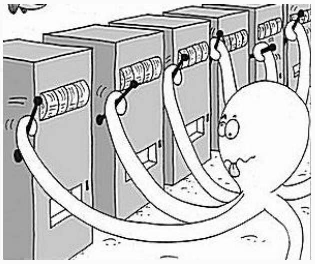


Image source : *Microsoft research*

- At each time step t, the agent selects an action $i(t)$ and then receives a numerical reward $r(t) \sim X_{i(t)}$ with mean $\mu_{i(t)}$.
- Agent's goal : Minimize the expected regret of its policy π

$$\mathfrak{R}_\pi(T) := \underbrace{T\mu_*}_{\text{Optimal expected cumulative reward}} - \underbrace{\mathbb{E}\left[\sum_{t=1}^{T} r(t) \mid \pi\right]}_{\text{Expected cumulative reward of } \pi}$$

where μ_* is the optimal mean reward and T is the horizon.

Recap Lecture 2 : Stationary stochastic bandits

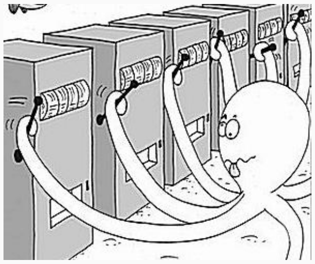


Image source : *Microsoft research*

- At each time step t, the agent selects an action $i(t)$ and then receives a numerical reward $r(t) \sim X_{i(t)}$ with mean $\mu_{i(t)}$.
- Agent's goal : Minimize the expected regret of its policy π

$$\mathfrak{R}_\pi(T) := \underbrace{T\mu_*}_{\text{Optimal expected cumulative reward}} - \underbrace{\mathbb{E}\left[\sum_{t=1}^{T} r(t) \mid \pi\right]}_{\text{Expected cumulative reward of } \pi}$$

where μ_* is the optimal mean reward and T is the horizon.

- Our aim : Construct an algorithm with sub-linear regret (featuring terms like $\sqrt{T}$ or $\log T$, but not T).

Recap Lecture 2 : UCB

Algorithm UCB algorithm Auer et al. [2002]

Parameters : Confidence level δ

for $t = 1, \ldots, K$ **do**

  Choose each arm once.

end for

for $t = K + 1, \ldots$ **do**

  Compute empirical means $\hat{\mu}_1(t-1), \ldots, \hat{\mu}_K(t-1)$.

  Select arm $i(t) = \arg\max_a \left[\hat{\mu}_a(t-1) + \sqrt{\frac{2\log(1/\delta)}{N_a(t-1)}}\right]$.

end for

Recap Lecture 2 : UCB

Algorithm UCB algorithm Auer et al. [2002]

Parameters : Confidence level δ

1: **for** $t = 1, \ldots, K$ **do**

2: Choose each arm once.

3: **end for**

4: **for** $t = K + 1, \ldots$ **do**

5: Compute empirical means $\hat{\mu}_1(t-1), \ldots, \hat{\mu}_K(t-1)$.

6: Select arm $i(t) = \arg\max_a \left[\hat{\mu}_a(t-1) + \sqrt{\frac{2\log(1/\delta)}{N_a(t-1)}}\right]$.

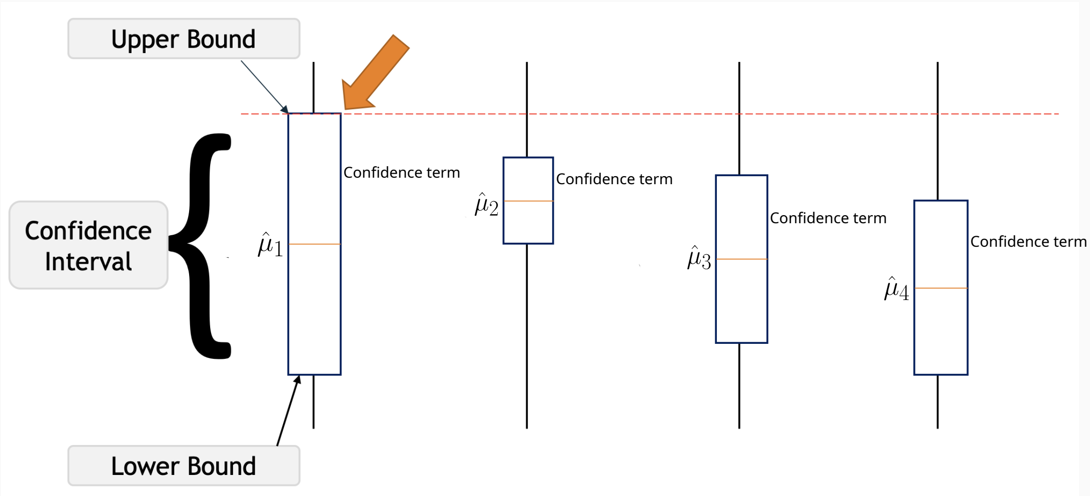
7: **end for**

- Distribution-dependent regret bound $\displaystyle\sum_{a:\Delta_a>0} \frac{16\log(T)}{\Delta_a} + 3\Delta_a$

  (recall that $\Delta_a = \mu_* - \mu_a$).
- Distribution-free regret bound $O(\sqrt{KT\log(T)})$.

$f(x) = O(g(x))$, if $f(x) < Cg(x)$ for all $x > n$. For more information, click here.

UCB : Solving Bandits from a Frequentist Perspective

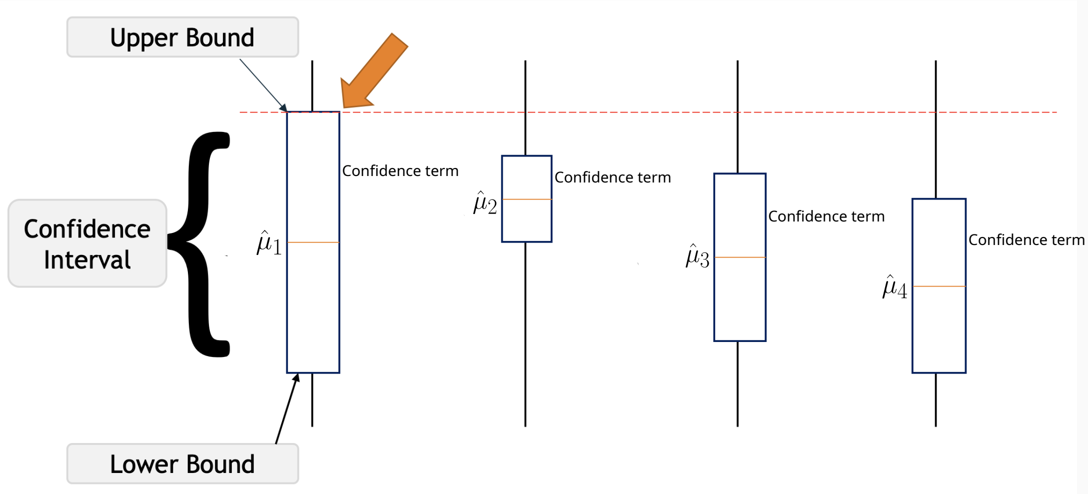


- Build confidence intervals around empirical mean rewards.

  Confidence term for arm $a = \sqrt{\frac{2\log(1/\delta)}{N_a(t-1)}}$

  Confidence interval for arm $a = \left\{\hat{\mu}_a - \sqrt{\frac{2\log(1/\delta)}{N_a(t-1)}}, \hat{\mu}_a + \sqrt{\frac{2\log(1/\delta)}{N_a(t-1)}}\right\}$

UCB : Solving Bandits from a Frequentist Perspective



- Build confidence intervals around empirical mean rewards.

  Confidence term for arm $a = \sqrt{\frac{2\log(1/\delta)}{N_a(t-1)}}$

  Confidence interval for arm $a = \left\{\hat{\mu}_a - \sqrt{\frac{2\log(1/\delta)}{N_a(t-1)}}, \hat{\mu}_a + \sqrt{\frac{2\log(1/\delta)}{N_a(t-1)}}\right\}$

- Arm selection rule using the size of the confidence interval.

  Select arm $i(t) = \arg\max_a \left[\hat{\mu}_a(t-1) + \sqrt{\frac{2\log(1/\delta)}{N_a(t-1)}}\right]$.

Lecture 3 : Outline

- Solving Bandits from a Bayesian Perspective
- Thompson Sampling
- Regret Bound for Thompson Sampling

Solving Bandits from a Bayesian Perspective

Solving Bandits from a Bayesian Perspective

Define a prior distribution that incorporates your subjective beliefs about unknown parameters i.e. mean rewards.

Solving Bandits from a Bayesian Perspective

Define a prior distribution that incorporates your subjective beliefs about unknown parameters i.e. mean rewards.

At each time step t,

1. Sample a particular set of parameters from the prior.

Solving Bandits from a Bayesian Perspective

Define a prior distribution that incorporates your subjective beliefs about unknown parameters i.e. mean rewards.
At each time step t,

1. Sample a particular set of parameters from the prior.
2. Select arm $i(t) = \arg\max_j \text{reward}_j \mid \text{parameters}$

Solving Bandits from a Bayesian Perspective

Define a prior distribution that incorporates your subjective beliefs about unknown parameters i.e. mean rewards.
At each time step t,

1. Sample a particular set of parameters from the prior.
2. Select arm $i(t) = \arg\max_j \text{reward}_j \mid \text{parameters}$
3. Observe reward and update posterior.

Solving Bandits from a Bayesian Perspective

Define a prior distribution that incorporates your subjective beliefs about unknown parameters i.e. mean rewards.
At each time step t,

1. Sample a particular set of parameters from the prior.
2. Select arm $i(t) = \arg\max_j \text{reward}_j \mid \text{parameters}$
3. Observe reward and update posterior.
   (Prior at time $t+1 \leftarrow$ posterior at time t)

Choice of Prior : Beta Prior

Solving bandits from a Bayesian perspective

Choose a prior for the mean reward of each arm.
At each time step,

1. Sample a particular set of parameters from the prior.
2. Select arm $i(t) = \arg\max_j \text{reward}_j \mid \text{parameters}$
3. Observe reward and update posterior.

Choice of Prior : Beta Prior

Solving bandits from a Bayesian perspective

Choose a prior for the mean reward of each arm.
At each time step,

1. Sample a particular set of parameters from the prior.
2. Select arm $i(t) = \arg\max_j \text{reward}_j \mid \text{parameters}$
3. Observe reward and update posterior.

- $Beta(\alpha, \beta)$ is a family of continuous distributions defined on $[0, 1]$.

Probability density function for $Beta(\alpha, \beta)$:

$$f(x, \alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\int_0^1 u^{\alpha-1}(1-u)^{\beta-1}du}$$

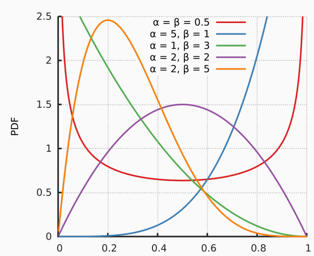
Choice of Prior : Beta Prior

Solving bandits from a Bayesian perspective

Choose a prior for the mean reward of each arm.
At each time step,

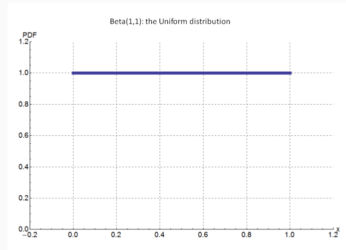
1. Sample a particular set of parameters from the prior.
2. Select arm $i(t) = \arg\max_j \text{reward}_j \mid \text{parameters}$
3. Observe reward and update posterior.

- $Beta(\alpha, \beta)$ is a family of continuous distributions defined on $[0, 1]$.



Probability density function for $Beta(\alpha, \beta)$:

$$f(x, \alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\int_0^1 u^{\alpha-1}(1-u)^{\beta-1} du}$$



- $Beta(1, 1) \equiv$ uniform distribution on $[0, 1]$.

Updating Posterior : Bernoulli Rewards using Beta Prior

Solving bandits from a Bayesian perspective

Choose a prior for the mean reward of each arm.
At each time step,

1. Sample a particular set of parameters from the prior.
2. Select arm $i(t) = \arg\max_j \text{reward} \mid \text{parameters}$
3. **Observe reward and update posterior**.

Updating Posterior : Bernoulli Rewards using Beta Prior

Solving bandits from a Bayesian perspective

Choose a prior for the mean reward of each arm.
At each time step,

1. Sample a particular set of parameters from the prior.
2. Select arm $i(t) = \arg\max_j$ reward | parameters
3. **Observe reward and update posterior**.

- For Bernoulli rewards (i.e. rewards either 0 or 1), interpret $Beta(\alpha, \beta)$ parameters as follows :
  - $\alpha - 1$ as the number of previous $1's$ and
  - $\beta - 1$ as the number of previous $0's$.

Updating Posterior : Bernoulli Rewards using Beta Prior

Solving bandits from a Bayesian perspective

Choose a prior for the mean reward of each arm.
At each time step,

1. Sample a particular set of parameters from the prior.
2. Select arm $i(t) = \arg\max_j$ reward | parameters
3. **Observe reward and update posterior**.

- For Bernoulli rewards (i.e. rewards either 0 or 1), interpret $\mathit{Beta}(\alpha, \beta)$ parameters as follows :
  - $\alpha - 1$ as the number of previous $1's$ and
  - $\beta - 1$ as the number of previous $0's$.
- After observing a Bernoulli reward,
  if the reward is 1,
  then the posterior distribution is $\mathit{Beta}(\alpha + 1, \beta)$
  if the reward is 0,
  then the posterior distribution is $\mathit{Beta}(\alpha, \beta + 1)$.

Updating Posterior: Bernoulli Rewards using Beta Prior

Solving bandits from a Bayesian perspective

Choose a prior for the mean reward of each arm.
At each time step,

1. Sample a particular set of parameters from the prior.
2. Select arm $i(t) = \arg\max_j$ reward | parameters
3. **Observe reward and update posterior**.

- For Bernoulli rewards (i.e. rewards either 0 or 1), interpret $Beta(\alpha, \beta)$ parameters as follows:
  - $\alpha - 1$ as the number of previous $1's$ and
  - $\beta - 1$ as the number of previous $0's$.
- After observing a Bernoulli reward,
  if the reward is 1,
  then the posterior distribution is $Beta(\alpha + 1, \beta)$
  if the reward is 0,
  then the posterior distribution is $Beta(\alpha, \beta + 1)$.

Why Beta prior? Because Beta is the conjugate prior for Bernoulli distribution. For more information, click here.

Thompson Sampling

Thompson Sampling algorithm

Algorithm Thompson sampling with Beta prior for Bernoulli rewards

for $i = 1, \ldots, K$ **do**
  Initialize $\text{Success}_i = 0$ and $\text{Failure}_i = 0$
end for
for $t = 1, \ldots, T$ **do**
  for $i = 1, \ldots, K$ **do**
    Sample $\theta_i(t) \sim \mathit{Beta}(\text{Success}_i + 1, \text{Failure}_i + 1)$
  end for
  Select arm $i(t) = \arg\max_j \theta_j(t)$.
  Observe reward $r(t)$.
  if $r(t) = 1$ **then**
    $\text{Success}_{i(t)} = \text{Success}_{i(t)} + 1$
  else
    $\text{Failure}_{i(t)} = \text{Failure}_{i(t)} + 1$
  end if
end for

Thompson Sampling algorithm

Algorithm Thompson sampling with Beta prior for Bernoulli rewards

```
for i = 1, ..., K do
    Initialize Success_i = 0 and Failure_i = 0
end for
for t = 1, ..., T do
    for i = 1, ..., K do
        Sample θ_i(t) ~ Beta(Success_i + 1, Failure_i + 1)
    end for
    Select arm i(t) = arg max_j θ_j(t).
    Observe reward r(t).
    if r(t) = 1 then
        Success_{i(t)} = Success_{i(t)} + 1
    else
        Failure_{i(t)} = Failure_{i(t)} + 1
    end if
end for
```

Thompson Sampling algorithm

Algorithm Thompson sampling with Beta prior for Bernoulli rewards

```
for i = 1, ..., K do
    Initialize Success_i = 0 and Failure_i = 0
end for
for t = 1, ..., T do
    for i = 1, ..., K do
        Sample θ_i(t) ~ Beta(Success_i + 1, Failure_i + 1)
    end for
    Select arm i(t) = arg max_j θ_j(t).
    Observe reward r(t).
    if r(t) = 1 then
        Success_{i(t)} = Success_{i(t)} + 1
    else
        Failure_{i(t)} = Failure_{i(t)} + 1
    end if
end for
```

Thompson Sampling algorithm

Algorithm Thompson sampling with Beta prior for Bernoulli rewards

for $i = 1, \ldots, K$ **do**
  Initialize $\text{Success}_i = 0$ and $\text{Failure}_i = 0$
end for
for $t = 1, \ldots, T$ **do**
  for $i = 1, \ldots, K$ **do**
    Sample $\theta_i(t) \sim \mathit{Beta}(\text{Success}_i + 1, \text{Failure}_i + 1)$
  end for
  Select arm $i(t) = \arg\max_j \theta_j(t)$.
  Observe reward $r(t)$.
  if $r(t) = 1$ **then**
    $\text{Success}_{i(t)} = \text{Success}_{i(t)} + 1$
  else
    $\text{Failure}_{i(t)} = \text{Failure}_{i(t)} + 1$
  end if
end for

Thompson Sampling algorithm

Algorithm Thompson sampling with Beta prior for Bernoulli rewards

```
for i = 1, ..., K do
    Initialize Success_i = 0 and Failure_i = 0
end for
for t = 1, ..., T do
    for i = 1, ..., K do
        Sample θ_i(t) ~ Beta(Success_i + 1, Failure_i + 1)
    end for
    Select arm i(t) = arg max_j θ_j(t).
    Observe reward r(t).
    if r(t) = 1 then
        Success_{i(t)} = Success_{i(t)} + 1
    else
        Failure_{i(t)} = Failure_{i(t)} + 1
    end if
end for
```

Thompson Sampling algorithm

Algorithm Thompson sampling with Beta prior for Bernoulli rewards

for $i = 1, \dots, K$ **do**
    Initialize $\text{Success}_i = 0$ and $\text{Failure}_i = 0$
end for
for $t = 1, \dots, T$ **do**
    for $i = 1, \dots, K$ **do**
        Sample $\theta_i(t) \sim \mathit{Beta}(\text{Success}_i + 1, \text{Failure}_i + 1)$
    end for
    Select arm $i(t) = \arg\max_j \theta_j(t)$.
    Observe reward $r(t)$.
    if $r(t) = 1$ **then**
        $\text{Success}_{i(t)} = \text{Success}_{i(t)} + 1$
    else
        $\text{Failure}_{i(t)} = \text{Failure}_{i(t)} + 1$
    end if
end for

Thompson Sampling algorithm

Algorithm Thompson sampling with Beta prior for Bernoulli rewards

```
for i = 1, ..., K do
    Initialize Success_i = 0 and Failure_i = 0
end for
for t = 1, ..., T do
    for i = 1, ..., K do
        Sample θ_i(t) ~ Beta(Success_i + 1, Failure_i + 1)
    end for
    Select arm i(t) = arg max_j θ_j(t).
    Observe reward r(t).
    if r(t) = 1 then
        Success_{i(t)} = Success_{i(t)} + 1
    else
        Failure_{i(t)} = Failure_{i(t)} + 1
    end if
end for
```

Why does Thompson Sampling work?

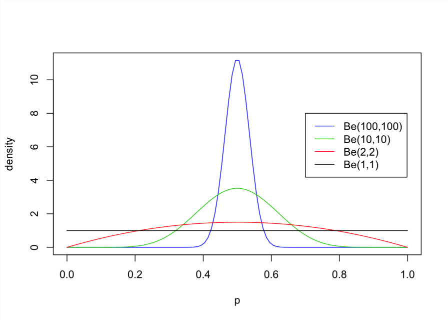
- Arm selection : Select arm $i(t) = \arg\max_j \theta_j(t)$.

Why does Thompson Sampling work?

- Arm selection : Select arm $i(t) = \arg\max_j \theta_j(t)$.
- Exploration via randomization
  $\theta_i(t) \sim \mathit{Beta}(\text{Success}_i + 1, \text{Failure}_i + 1)$

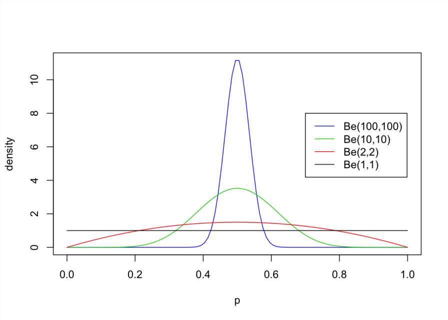
Why does Thompson Sampling work?

- Arm selection : Select arm $i(t) = \arg\max_j \theta_j(t)$.
- Exploration via randomization
  $\theta_i(t) \sim Beta(\text{Success}_i + 1, \text{Failure}_i + 1)$
- Initially the posterior might be poorly concentrated, then the fluctuations in θ's are likely to be large and TS will explore.

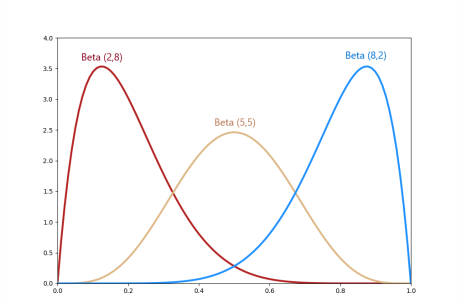


Why does Thompson Sampling work?

- Arm selection : Select arm $i(t) = \arg\max_j \theta_j(t)$.
- Exploration via randomization
  $\theta_i(t) \sim \mathit{Beta}(\text{Success}_i + 1, \text{Failure}_i + 1)$
- Initially the posterior might be poorly concentrated, then the fluctuations in θ's are likely to be large and TS will explore.



- After a large number of observations, the posterior concentrates around the true mean and the rate of exploration decreases.



Regret Bound for Thompson Sampling

Regret Bound for Thompson Sampling

Theorem (Theorem 1 from Agrawal and Goyal [2013])

After T time steps, the expected cumulative regret of Thompson sampling using Beta priors is

$$Regret = \mathfrak{R}(T) \leq (1+\epsilon)^2 \sum_i \frac{\log T}{c} \Delta_i + O\left(\frac{K}{\epsilon^2}\right),$$

where c is a problem-dependent constant.

Proving the Regret Bound : Preliminaries I

- True mean reward of arm i is μ_i.
- By default, 1 is the optimal arm i.e. μ_1 is the optimal mean.

Proving the Regret Bound : Preliminaries I

- True mean reward of arm i is μ_i.
- By default, 1 is the optimal arm i.e. μ_1 is the optimal mean.
- Arm being played at time $t = i(t)$.

Proving the Regret Bound : Preliminaries I

- True mean reward of arm i is μ_i.
- By default, 1 is the optimal arm i.e. μ_1 is the optimal mean.
- Arm being played at time $t = i(t)$.
- $N_i(t) :=$ Number of times arm i is played till $t = \sum_{\tau=1}^{t} \mathbb{I}(i(\tau) = i)$.

Proving the Regret Bound : Preliminaries I

- True mean reward of arm i is μ_i.
- By default, 1 is the optimal arm i.e. μ_1 is the optimal mean.
- Arm being played at time $t = i(t)$.
- $N_i(t) :=$ Number of times arm i is played till $t = \sum_{\tau=1}^{t} \mathbb{I}(i(\tau) = i)$.
- Empirical mean of arm i at $t = \hat{\mu}_i(t) := \frac{1}{N_i(t)} \sum_{\tau=1}^{t} (r(\tau) | i(\tau) = i)$.

Proving the Regret Bound : Preliminaries I

- True mean reward of arm i is μ_i.
- By default, 1 is the optimal arm i.e. μ_1 is the optimal mean.
- Arm being played at time $t = i(t)$.
- $N_i(t) :=$ Number of times arm i is played till $t = \sum_{\tau=1}^{t} \mathbb{I}(i(\tau) = i)$.
- Empirical mean of arm i at $t = \hat{\mu}_i(t) := \frac{1}{N_i(t)} \sum_{\tau=1}^{t} (r(\tau)|i(\tau) = i)$.
- Sampled parameter of arm $i = \theta_i(t)$.

Proving the Regret Bound : Preliminaries II

Recall from the last lecture

Lemma

$$\textit{Regret} = \mathfrak{R}(T) = \sum_{i=1,\ldots,K,\Delta_i>0} \Delta_i \, \mathbb{E}[N_i(T)].$$

- Suboptimality gap $\Delta_i := \mu_* - \mu_i$ where μ_* is the optimal mean reward and μ_i is the mean reward for arm a.
- $N_i(T) :=$ Number of times arm i is played till $T = \sum_{t=1}^{T} \mathbb{I}(i(t) = i)$.

Proving the Regret Bound : Preliminaries II

Recall from the last lecture

Lemma

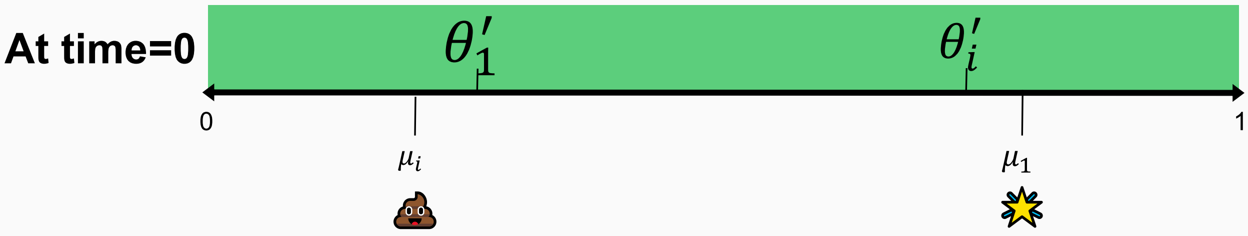
$$Regret = \mathfrak{R}(T) = \sum_{i=1,\ldots,K,\Delta_i>0} \Delta_i \, \mathbb{E}[N_i(T)].$$

- Suboptimality gap $\Delta_i := \mu_* - \mu_i$ where μ_* is the optimal mean reward and μ_i is the mean reward for arm a.
- $N_i(T) :=$ Number of times arm i is played till $T = \sum_{t=1}^{T} \mathbb{I}(i(t) = i)$.
- In order to bound $\mathfrak{R}(T)$, we need to bound $\mathbb{E}[N_i(T)]$.

When does Thompson Sampling Perform Well? I

Arm selection rule of Thompson sampling

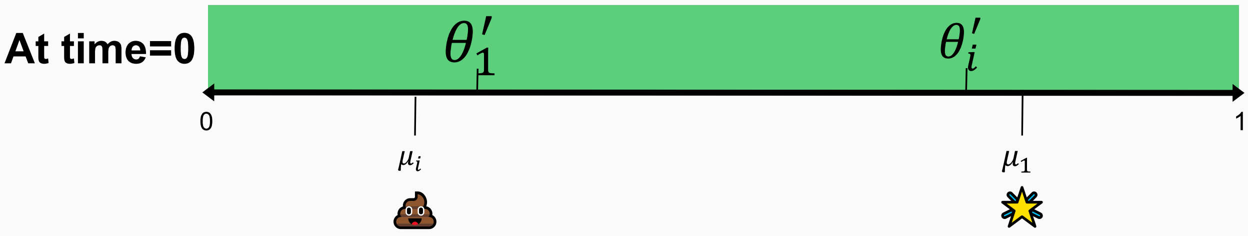
Select arm $i(t) = \arg\max_j \theta_j(t)$.



When does Thompson Sampling Perform Well? I

Arm selection rule of Thompson sampling

Select arm $i(t) = \arg\max_j \theta_j(t)$.

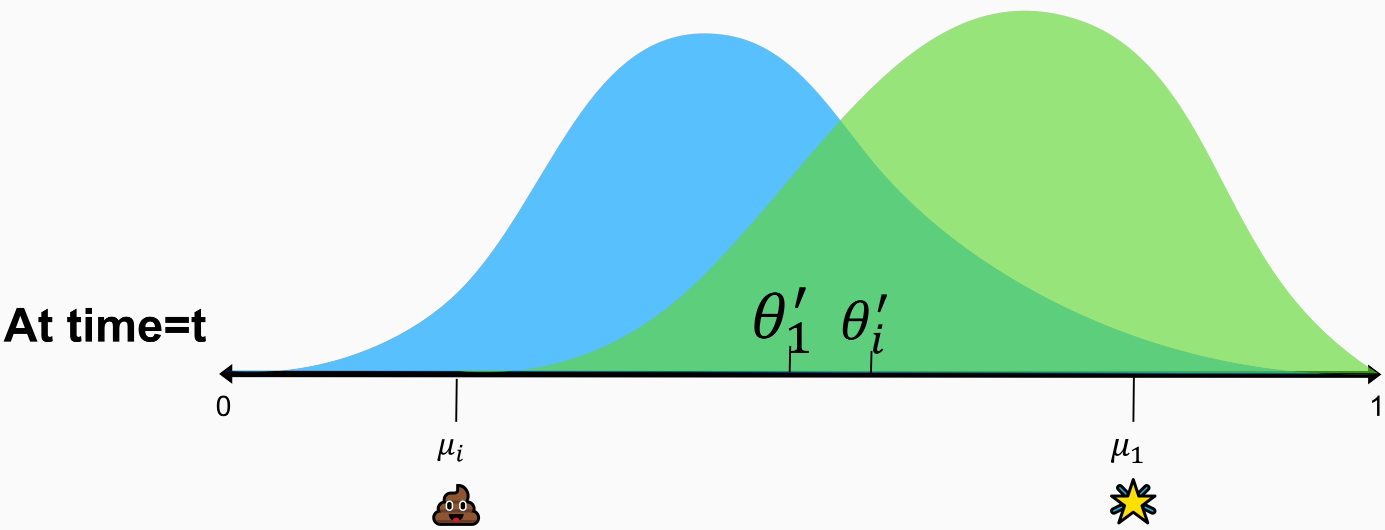


- Initially, all θ's are from the same distribution *Beta*$(1,1)$ (i.e., the uniform distribution on $[0,1]$), so not yet! ☹

When does Thompson Sampling Perform Well? II

Arm selection rule of Thompson sampling

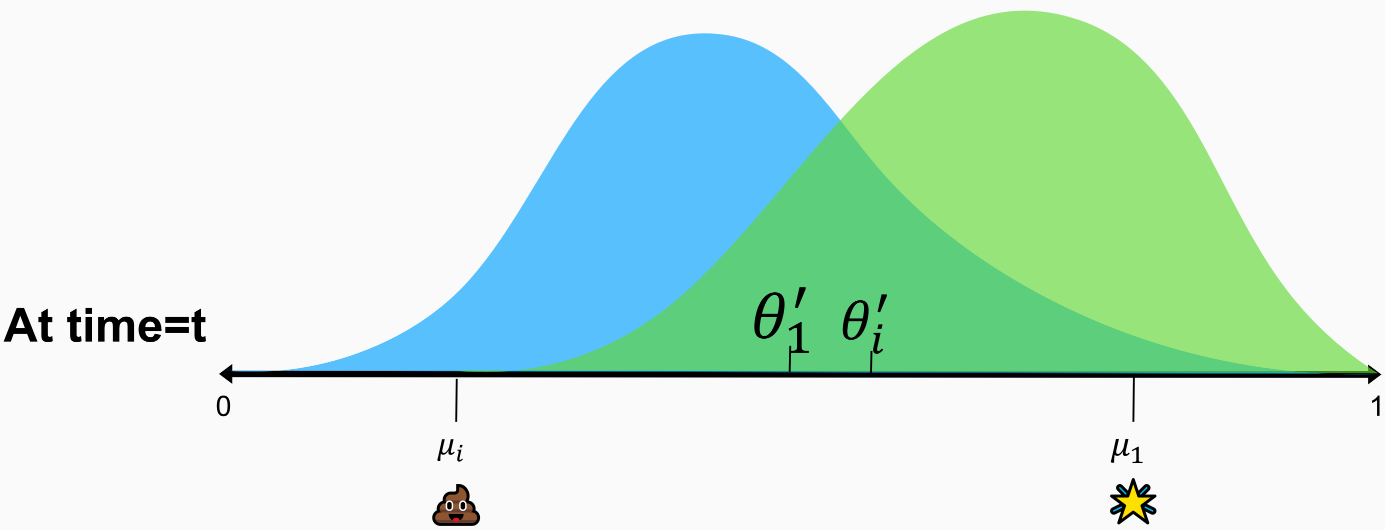
Select arm $i(t) = \arg\max_j \theta_j(t)$.



When does Thompson Sampling Perform Well? II

Arm selection rule of Thompson sampling

Select arm $i(t) = \arg\max_j \theta_j(t)$.

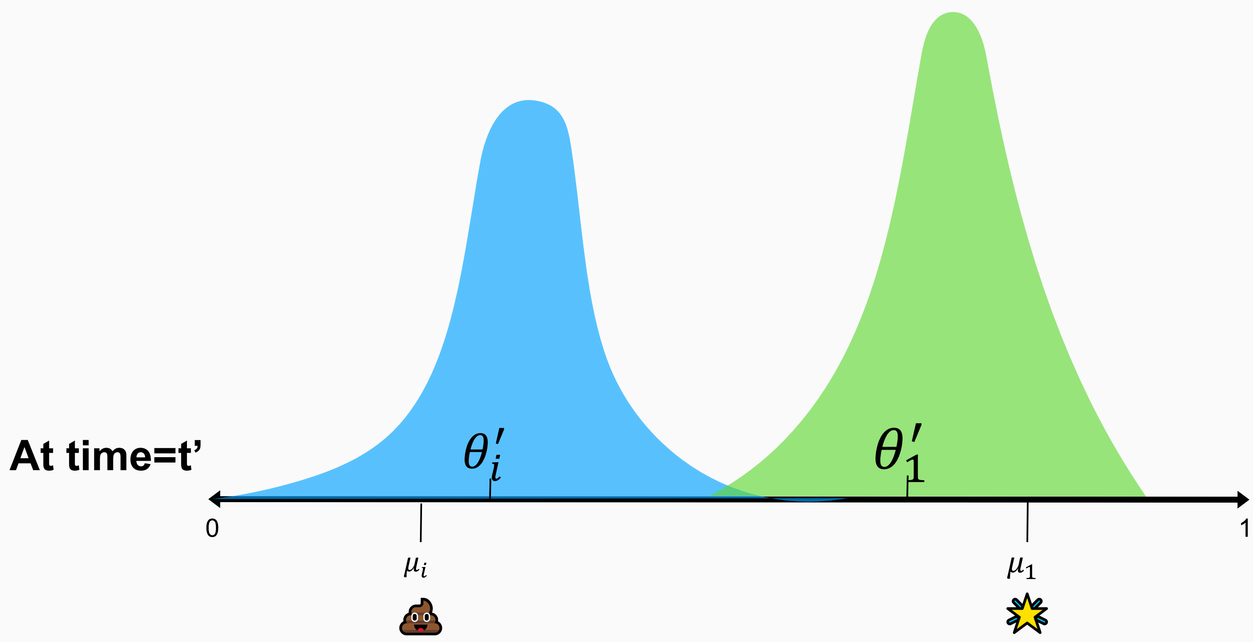


- At t, θ's are too far from μ's, so not yet! ☹

When does Thompson Sampling Perform Well? III

Arm selection rule of Thompson sampling

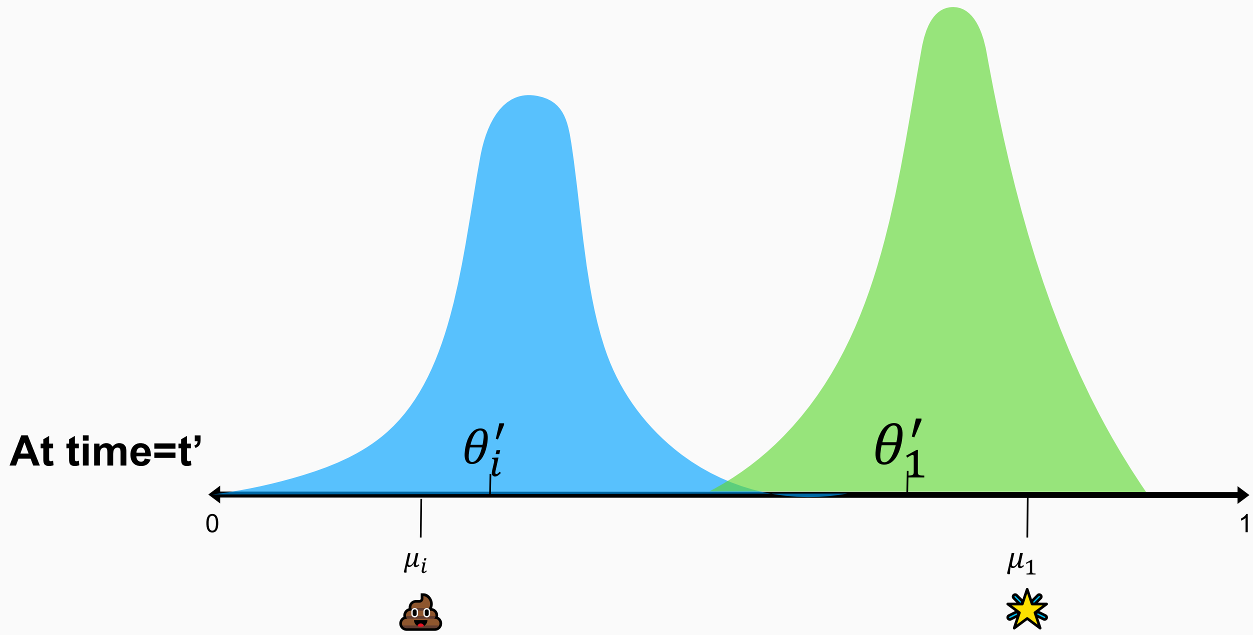
Select arm $i(t) = \arg\max_j \theta_j(t)$.



When does Thompson Sampling Perform Well? III

Arm selection rule of Thompson sampling

Select arm $i(t) = \arg\max_j \theta_j(t)$.



- At t', when θ's are close μ's. 🙂

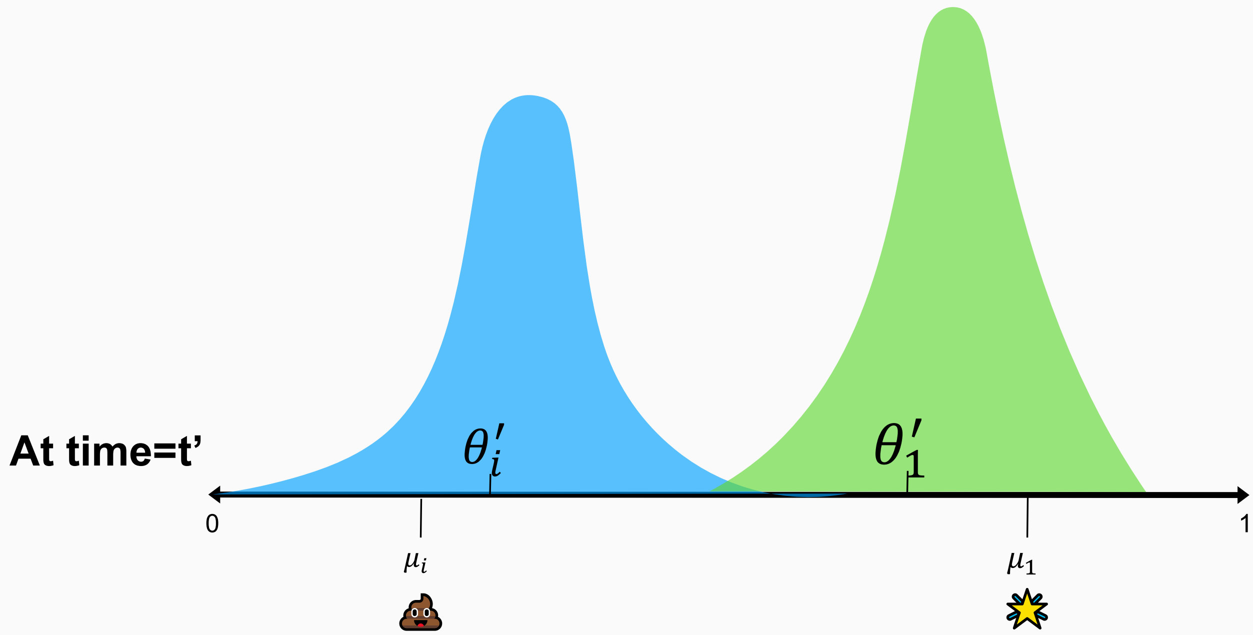
Break

We start again after a break.

When does Thompson Sampling Perform Well?

Arm selection rule of Thompson sampling

Select arm $i(t) = \arg\max_j \theta_j(t)$.

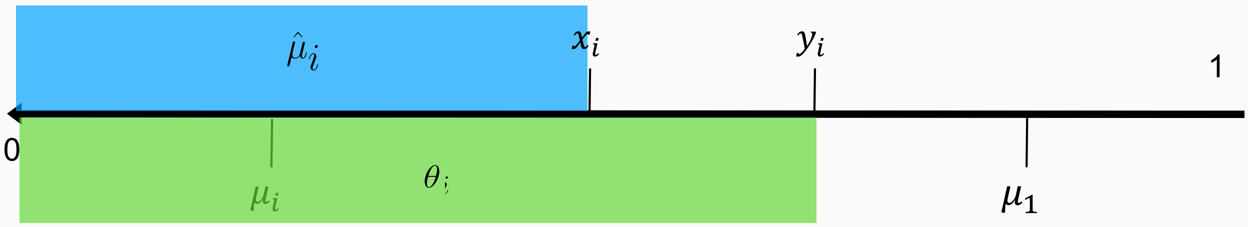


- At t', when θ's are close μ's. 🙂

Proving the Regret Bound : Defining the Good Events

- $E_i^{\theta}(t) :=$ sampled parameter θ_i is close to μ_i.
- $E_i^{\mu}(t) :=$ estimated mean $\hat{\mu}_i$ is close to μ_i.

Proving the Regret Bound : Defining the Good Events



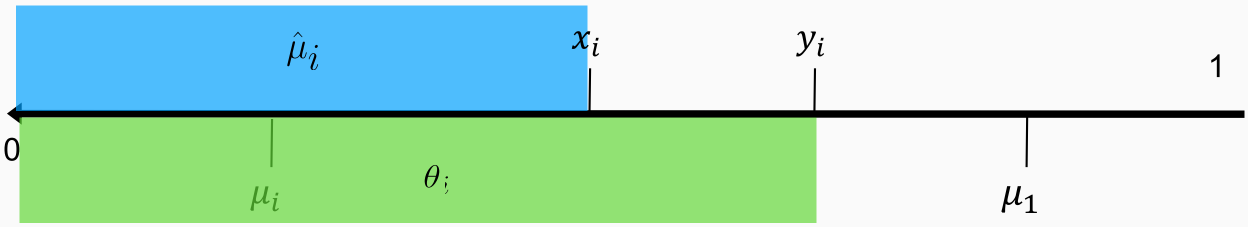
- For each suboptimal arm i, let x_i and y_i be two thresholds such that $\mu_i < x_i < y_i < \mu_1$.
- $E_i^{\theta}(t) :=$ sampled parameter θ_i is close to μ_i,

$$\boxed{E_i^{\theta}(t) := \{\theta_i < y_i\}}.$$

- $E_i^{\mu}(t) :=$ estimated mean $\hat{\mu}_i$ is close to μ_i,

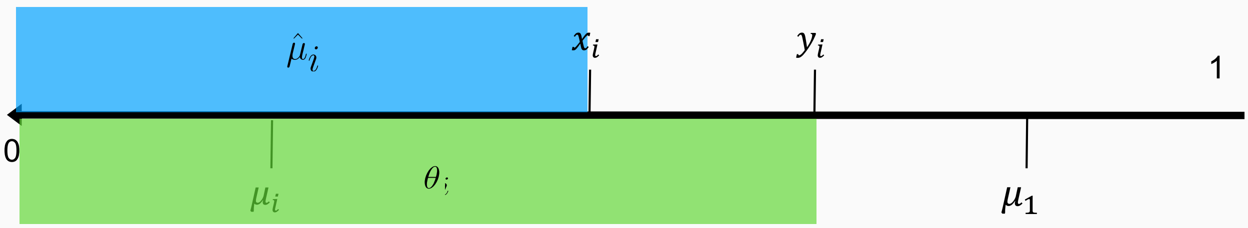
$$\boxed{E_i^{\mu}(t) := \{\hat{\mu}_i < x_i\}}.$$

Proving the Regret Bound : Decomposition into Three Terms



$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}(i(t) = i)$$

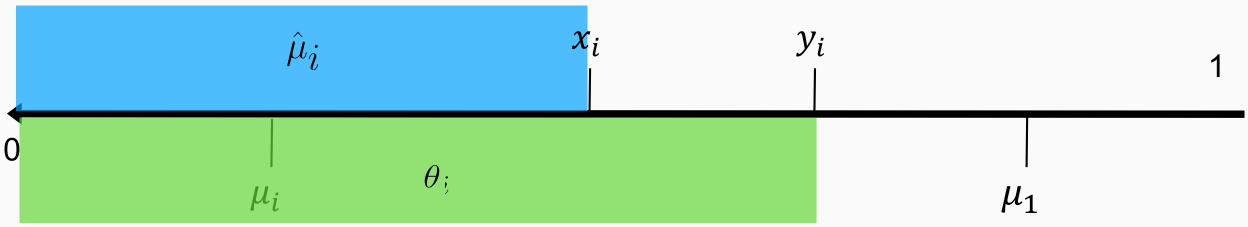
Proving the Regret Bound : Decomposition into Three Terms



$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}(i(t) = i)$$

$$= \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)} \right) + \dots$$

Proving the Regret Bound : Decomposition into Three Terms

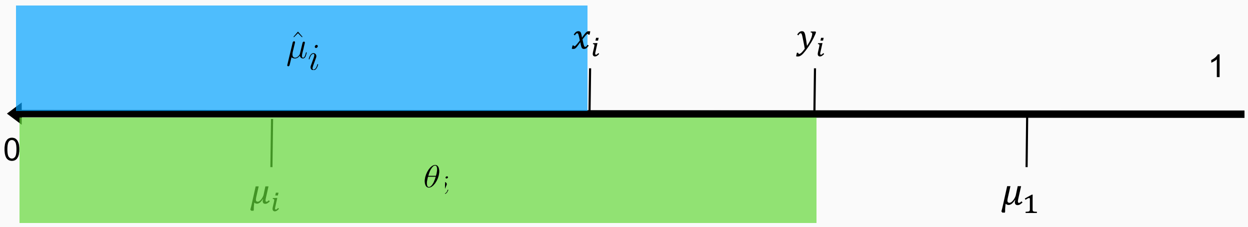


$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}(i(t) = i)$$

$$= \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, E_i^{\mu}(t), E_i^{\theta}(t)\right)$$

$$+ \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, E_i^{\mu}(t), \overline{E_i^{\theta}(t)}\right) + \dots$$

Proving the Regret Bound : Decomposition into Three Terms



$$
\begin{aligned}
\mathbb{E}[N_i(T)] &= \sum_{t=1}^{T} \mathbb{P}(i(t) = i) \\
&= \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, E_i^{\mu}(t), E_i^{\theta}(t)\right) \\
&\quad + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, E_i^{\mu}(t), \overline{E_i^{\theta}(t)}\right) \\
&\quad + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \overline{E_i^{\mu}(t)}\right)
\end{aligned}
$$

Proving the Regret Bound : Analyzing the First Term I

$$\mathbb{E}[N_i(T)] = \boxed{\sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)}\right)} + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \overline{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \overline{E_i^{\mu}(t)}\right)$$

Proving the Regret Bound: Analyzing the First Term I

$$\mathbb{E}[N_i(T)] = \boxed{\sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)}\right)} + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \overline{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \overline{E_i^{\mu}(t)}\right)$$

- Let "history" $\mathcal{F}_{t-1} = i(1), r(1), i(2), r(2), \ldots, i(t-1), r(t-1)$.

Proving the Regret Bound: Analyzing the First Term I

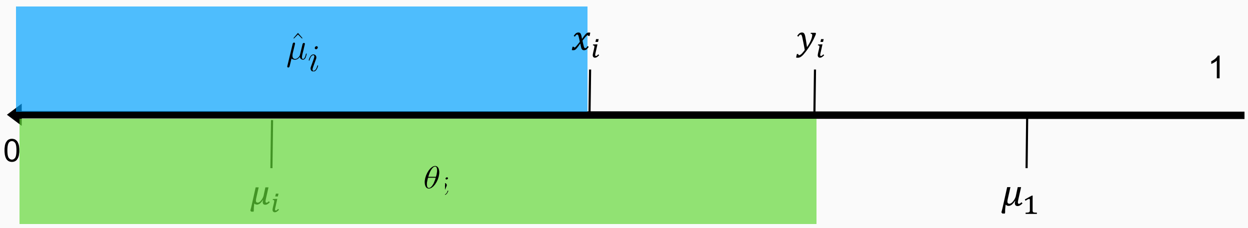
$$\mathbb{E}[N_i(T)] = \boxed{\sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)}\right)} + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \overline{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \overline{E_i^{\mu}(t)}\right)$$

- Let "history" $\mathcal{F}_{t-1} = i(1), r(1), i(2), r(2), \ldots, i(t-1), r(t-1)$.

Lemma (Main Lemma. Lemma 1 from Agrawal and Goyal [2013])

For all $t = 1, \ldots, T$ and all suboptimal arms i i.e. $i \neq 1$,

$$\mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)} \mid \mathcal{F}_{t-1}\right)$$
$$\leq \textit{Coefficient} \cdot \mathbb{P}\left(i(t) = 1, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)} \mid \mathcal{F}_{t-1}\right)$$



Proving the Regret Bound: Analyzing the First Term I

$$\mathbb{E}[N_i(T)] = \boxed{\sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)}\right)} + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \overline{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \overline{E_i^{\mu}(t)}\right)$$

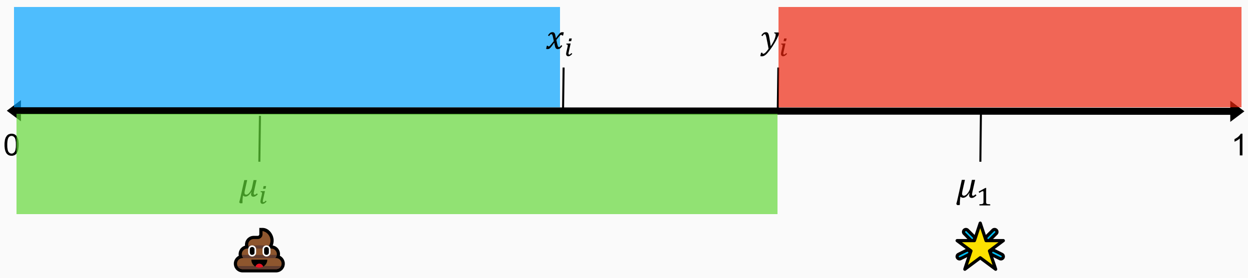
- Let "history" $\mathcal{F}_{t-1} = i(1), r(1), i(2), r(2), \ldots, i(t-1), r(t-1)$.

Lemma (Main Lemma. Lemma 1 from Agrawal and Goyal [2013])

For all $t = 1, \ldots, T$ and all suboptimal arms i i.e. $i \neq 1$,

$$\mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)} \mid \mathcal{F}_{t-1}\right) \leq \frac{(1 - p_{i,t})}{p_{i,t}} \mathbb{P}\left(i(t) = 1, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)} \mid \mathcal{F}_{t-1}\right)$$

where $p_{i,t} := \mathbb{P}(\boxed{\theta_1(t) > y_i} \mid \mathcal{F}_{t-1})$



Proving the Regret Bound : Analyzing the First Term II

$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ E_i^{\mu}(t),\ E_i^{\theta}(t) \right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ E_i^{\mu}(t),\ \overline{E_i^{\theta}(t)} \right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ \overline{E_i^{\mu}(t)} \right)$$

Proving the Regret Bound : Analyzing the First Term II

$$\mathbb{E}[N_i(T)] = \boxed{\sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \; E_i^{\mu}(t), \; E_i^{\theta}(t)\right)} + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \; E_i^{\mu}(t), \overline{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \overline{E_i^{\mu}(t)}\right)$$

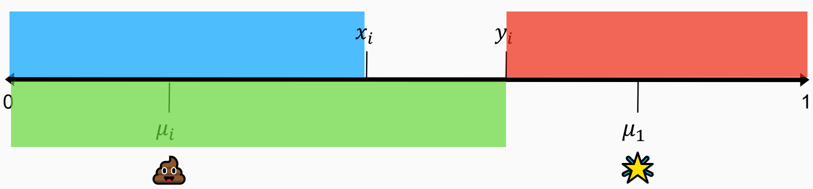
$$\text{First term} \leq \sum_{t=1}^{T} \mathbb{E}\left[\underbrace{\frac{(1 - p_{i,t})}{p_{i,t}}}_{\textit{Coefficient}} \cdot \underbrace{\mathbb{P}\left(i(t) = 1, \; E_i^{\mu}(t), \; E_i^{\theta}(t) \mid \mathcal{F}_{t-1}\right)}_{\text{Probability of playing the best arm in the "good" case}}\right]$$

Proving the Regret Bound : Analyzing the First Term II

$$\mathbb{E}[N_i(T)] = \boxed{\sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ E_i^{\mu}(t),\ E_i^{\theta}(t)\right)} + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ E_i^{\mu}(t),\ \overline{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ \overline{E_i^{\mu}(t)}\right)$$

$$\text{First term} \leq \sum_{t=1}^{T} \mathbb{E}\left[\underbrace{\frac{(1-p_{i,t})}{p_{i,t}}}_{\textit{Coefficient}} \cdot \underbrace{\mathbb{P}\left(i(t) = 1,\ E_i^{\mu}(t),\ E_i^{\theta}(t)\ \middle|\ \mathcal{F}_{t-1}\right)}_{\text{Probability of playing the best arm in the "good" case}}\right]$$

$$p_{i,t} := \mathbb{P}(\theta_1(t) > y_i \mid \mathcal{F}_{t-1})$$

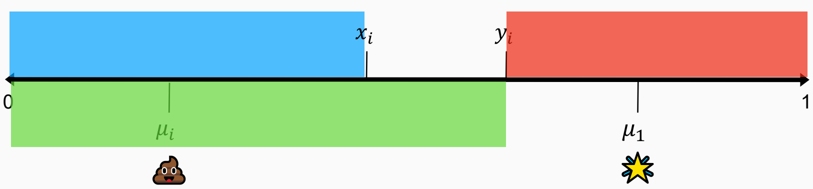


Proving the Regret Bound : Analyzing the First Term II

$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, E_i^{\mu}(t), E_i^{\theta}(t)\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, E_i^{\mu}(t), \overline{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \overline{E_i^{\mu}(t)}\right)$$

$$\text{First term} \leq \sum_{t=1}^{T} \mathbb{E}\left[\underbrace{\frac{(1-p_{i,t})}{p_{i,t}}}_{\textit{Coefficient}} \cdot \underbrace{\mathbb{P}\left(i(t) = 1, E_i^{\mu}(t), E_i^{\theta}(t) \mid \mathcal{F}_{t-1}\right)}_{\text{Probability of playing the best arm in the "good" case}}\right]$$

$$p_{i,t} := \mathbb{P}(\theta_1(t) > y_i \mid \mathcal{F}_{t-1})$$



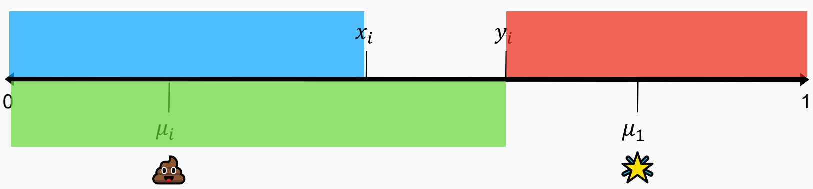
Coefficient decreases exponentially fast with samples of the optimal arm $N_1(t)$.

Proving the Regret Bound : Analyzing the First Term II

$$\mathbb{E}[N_i(T)] = \boxed{\sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \; E_i^{\mu}(t), \; E_i^{\theta}(t)\right)} + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \; E_i^{\mu}(t), \; \overline{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \; \overline{E_i^{\mu}(t)}\right)$$

$$\text{First term} \leq \sum_{t=1}^{T} \mathbb{E}\left[\underbrace{\frac{(1-p_{i,t})}{p_{i,t}}}_{\textit{Coefficient}} \cdot \underbrace{\mathbb{P}\left(i(t) = 1, \; E_i^{\mu}(t), \; E_i^{\theta}(t) \; | \; \mathcal{F}_{t-1}\right)}_{\text{Probability of playing the best arm in the "good" case}} \right]$$

$$p_{i,t} := \mathbb{P}(\theta_1(t) > y_i \, | \mathcal{F}_{t-1})$$



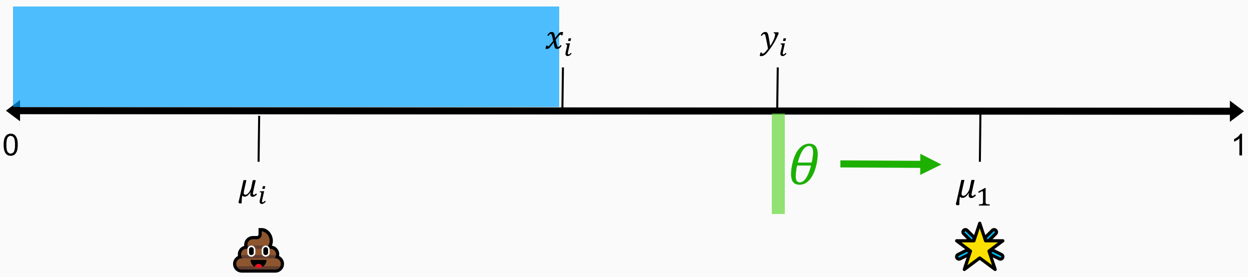
Coefficient decreases exponentially fast with samples of the optimal arm $N_1(t)$.

The term $\sum_{t=1}^{T} \mathbb{P}(i(t) = i, \; E_i^{\mu}(t), \; E_i^{\theta}(t))$ contributes a constant $O(1)$.

A primer on *big-oh* notation $O()$

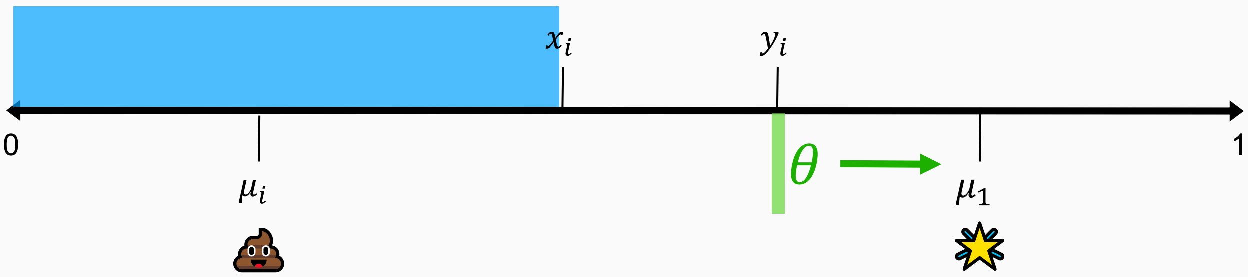
Proving the Regret Bound : Analyzing the Second Term I

$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t)=i,\ E_i^{\mu}(t),\ E_i^{\theta}(t)\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t)=i,\ E_i^{\mu}(t),\ \overline{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t)=i,\ \overline{E_i^{\mu}(t)}\right)$$



Proving the Regret Bound : Analyzing the Second Term I

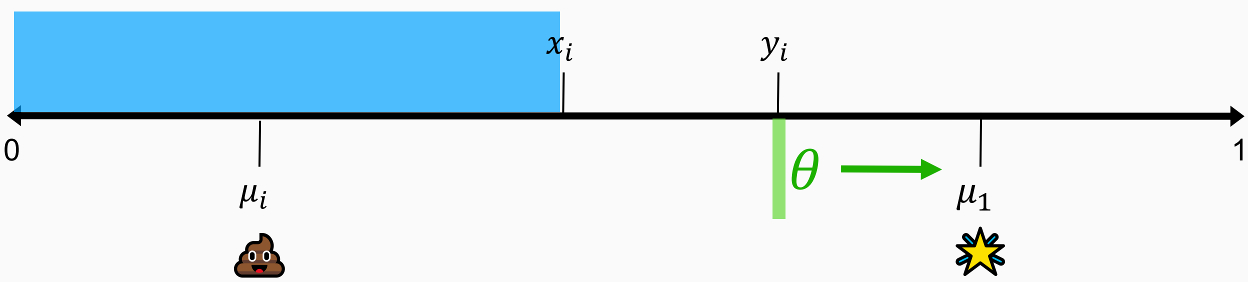
$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)}\right) + \boxed{\sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \boxed{\overline{E_i^{\theta}(t)}}\right)} + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \overline{E_i^{\mu}(t)}\right)$$



Proof sketch.

Proving the Regret Bound: Analyzing the Second Term I

$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ E_i^{\mu}(t),\ E_i^{\theta}(t)\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ E_i^{\mu}(t),\ \overline{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ \overline{E_i^{\mu}(t)}\right)$$

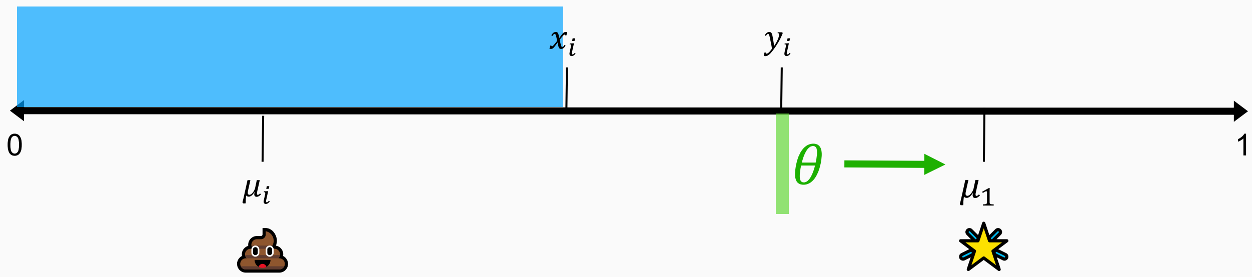


Proof sketch.

- Given that $E_i^{\mu}(t)$ holds, i.e. $\hat{\mu}_i \leq x_i$, the algorithm can only sample $\theta_i > y_i$ before the posterior concentrates around its mean.

Proving the Regret Bound : Analyzing the Second Term I

$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ \boxed{E_i^{\mu}(t)},\ \boxed{E_i^{\theta}(t)}\right) + \boxed{\sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ \boxed{E_i^{\mu}(t)},\ \boxed{\overline{E_i^{\theta}(t)}}\right)} + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ \overline{E_i^{\mu}(t)}\right)$$



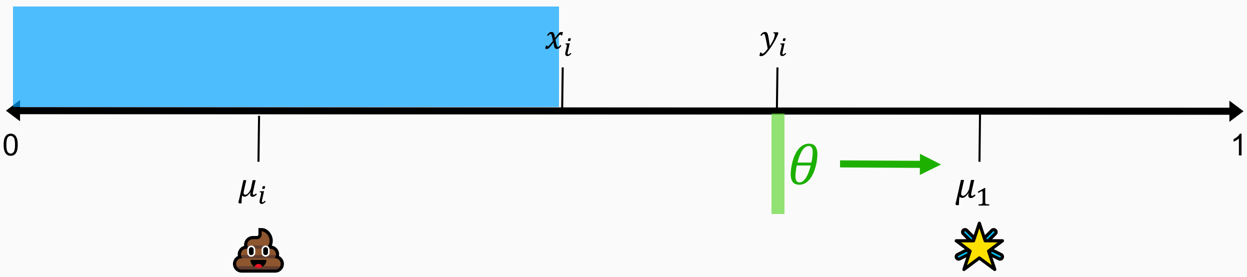
Proof sketch.

- Given that $\boxed{E_i^{\mu}(t)}$ holds, i.e. $\hat{\mu}_i \leq x_i$, the algorithm can only sample $\boxed{\theta_i > y_i}$ before the posterior concentrates around its mean.
- Posterior is well-concentrated around its mean when $N_i(t) \geq \frac{\log T}{d(x_i, y_i)}$,

$$d(x_i, y_i) := x_i \log \frac{x_i}{y_i} + (1 - x_i) \log \frac{1 - x_i}{1 - y_i}$$

Proving the Regret Bound : Analyzing the Second Term I

$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ E_i^{\mu}(t),\ E_i^{\theta}(t)\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ E_i^{\mu}(t),\ \overline{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ \overline{E_i^{\mu}(t)}\right)$$



Proof sketch.

- Given that $E_i^{\mu}(t)$ holds, i.e. $\hat{\mu}_i \leq x_i$, the algorithm can only sample $\theta_i > y_i$ before the posterior concentrates around its mean.
- Posterior is well-concentrated around its mean when $N_i(t) \geq \frac{\log T}{d(x_i, y_i)}$,

$$d(x_i, y_i) := x_i \log \frac{x_i}{y_i} + (1 - x_i) \log \frac{1 - x_i}{1 - y_i}$$

- After that, $\mathbb{P}(\theta_i > y_i)$ i.e. $\mathbb{P}\left(\overline{E_i^{\theta}(t)}\right) \leq \frac{1}{T}$.

Proving the Regret Bound : Analyzing the Second Term II

$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, E_i^{\mu}(t), E_i^{\theta}(t)\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, E_i^{\mu}(t), \overline{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \overline{E_i^{\mu}(t)}\right)$$

Proving the Regret Bound : Analyzing the Second Term II

$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)} \right) + \boxed{\sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \overline{E_i^{\theta}(t)} \right)} + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \overline{E_i^{\mu}(t)} \right)$$

- After $N_i(t) > \frac{\log T}{d(x_i, y_i)}$, $\mathbb{P}\left(\boxed{E_i^{\mu}(t)}, \overline{E_i^{\theta}(t)} \right) \le \frac{1}{T}$.

Proving the Regret Bound : Analyzing the Second Term II

$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \; E_i^{\mu}(t) \; , \; E_i^{\theta}(t)\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \; E_i^{\mu}(t) \; , \overline{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \overline{E_i^{\mu}(t)}\right)$$

- After $N_i(t) > \frac{\log T}{d(x_i, y_i)}$, $\mathbb{P}\left(E_i^{\mu}(t) , \overline{E_i^{\theta}(t)}\right) \leq \frac{1}{T}$.
- Note that $\mathbb{P}(\text{event}) = \mathbb{E}[\mathbb{I}(\text{event})]$.

Proving the Regret Bound: Analyzing the Second Term II

$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)}\right) + \boxed{\sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \overline{E_i^{\theta}(t)}\right)} + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \overline{E_i^{\mu}(t)}\right)$$

- After $N_i(t) > \frac{\log T}{d(x_i, y_i)}$, $\mathbb{P}\left(\boxed{E_i^{\mu}(t)}, \overline{E_i^{\theta}(t)}\right) \leq \frac{1}{T}$.
- Note that $\mathbb{P}(\text{event}) = \mathbb{E}[\mathbb{I}(\text{event})]$.

$$\text{Second term} \leq \mathbb{E}\left[\sum_{t=1}^{T} \mathbb{I}\left(i(t) = i, N_i(t) \leq \frac{\log T}{d(x_i, y_i)}\right)\right] + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \overline{E_i^{\theta}(t)}, N_i(t) > \frac{\log T}{d(x_i, y_i)}\right)$$

Proving the Regret Bound : Analyzing the Second Term II

$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ \boxed{E_i^{\mu}(t)},\ \boxed{E_i^{\theta}(t)}\right) + \boxed{\sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ \boxed{E_i^{\mu}(t)},\ \overline{E_i^{\theta}(t)}\right)} + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ \overline{E_i^{\mu}(t)}\right)$$

- After $N_i(t) > \frac{\log T}{d(x_i, y_i)}$, $\mathbb{P}\left(\boxed{E_i^{\mu}(t)}, \overline{E_i^{\theta}(t)}\right) \leq \frac{1}{T}$.
- Note that $\mathbb{P}(\text{event}) = \mathbb{E}[\mathbb{I}(\text{event})]$.

$$\begin{aligned}
\text{Second term} &\leq \mathbb{E}\left[\sum_{t=1}^{T} \mathbb{I}\left(i(t) = i, N_i(t) \leq \frac{\log T}{d(x_i, y_i)}\right)\right] \\
&\quad + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \overline{E_i^{\theta}(t)}, N_i(t) > \frac{\log T}{d(x_i, y_i)}\right) \\
&\leq \mathbb{E}\left[\sum_{t=1}^{T} \mathbb{I}\left(i(t) = i, N_i(t) \leq \frac{\log T}{d(x_i, y_i)}\right)\right] + \sum_{t=1}^{T} \frac{1}{T}
\end{aligned}$$

Proving the Regret Bound : Analyzing the Second Term II

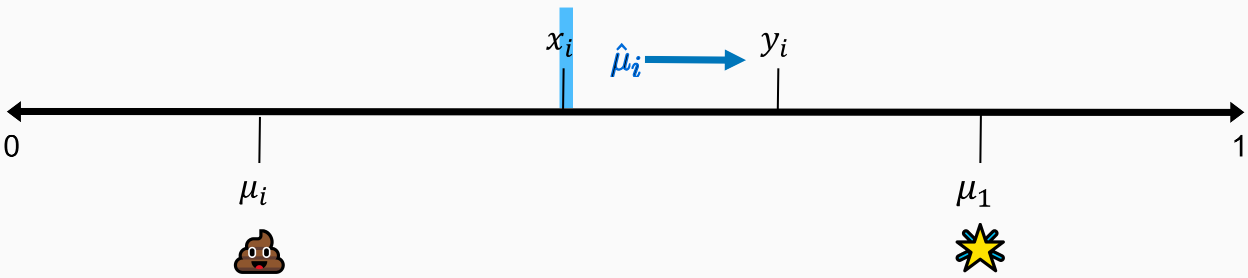
$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \; E_i^{\mu}(t), \; E_i^{\theta}(t)\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \; E_i^{\mu}(t), \; \overline{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \overline{E_i^{\mu}(t)}\right)$$

- After $N_i(t) > \frac{\log T}{d(x_i, y_i)}$, $\mathbb{P}\left(E_i^{\mu}(t), \overline{E_i^{\theta}(t)}\right) \leq \frac{1}{T}$.
- Note that $\mathbb{P}(\text{event}) = \mathbb{E}[\mathbb{I}(\text{event})]$.

$$\begin{aligned}
\text{Second term} &\leq \mathbb{E}\left[\sum_{t=1}^{T} \mathbb{I}\left(i(t) = i, N_i(t) \leq \frac{\log T}{d(x_i, y_i)}\right)\right] \\
&\quad + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \; E_i^{\mu}(t), \overline{E_i^{\theta}(t)}, N_i(t) > \frac{\log T}{d(x_i, y_i)}\right) \\
&\leq \mathbb{E}\left[\sum_{t=1}^{T} \mathbb{I}\left(i(t) = i, N_i(t) \leq \frac{\log T}{d(x_i, y_i)}\right)\right] + \sum_{t=1}^{T} \frac{1}{T} \\
&\leq \frac{\log T}{d(x_i, y_i)} + 1
\end{aligned}$$

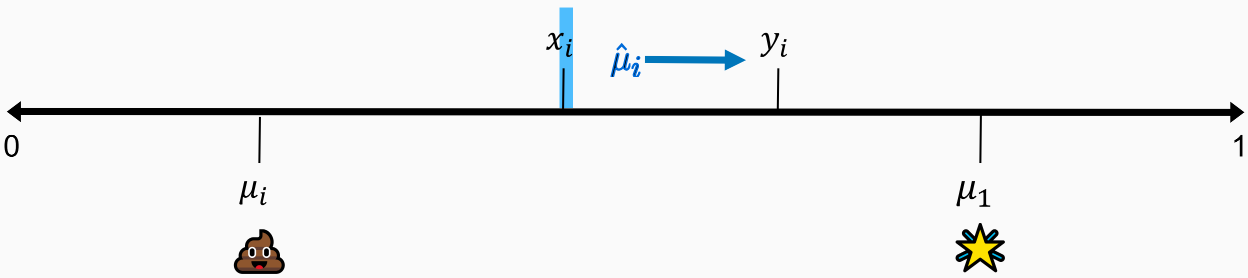
Proving the Regret Bound : Analyzing the Third Term

$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ E_i^{\mu}(t),\ E_i^{\theta}(t)\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ E_i^{\mu}(t),\ \overline{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ \overline{E_i^{\mu}(t)}\right)$$



Proving the Regret Bound : Analyzing the Third Term

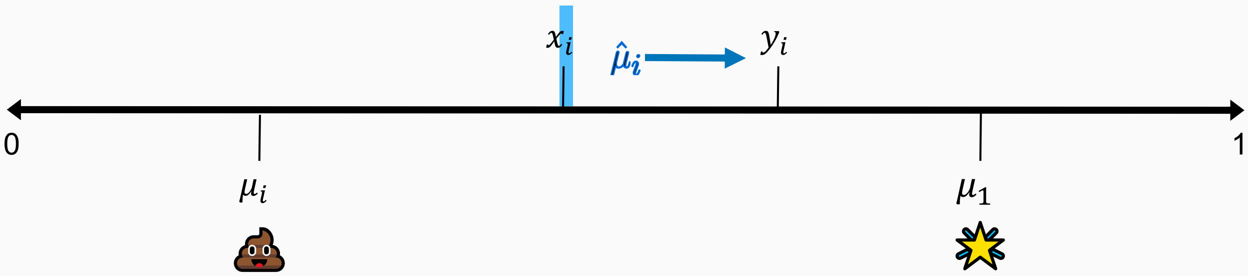
$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ E_i^{\mu}(t),\ E_i^{\theta}(t)\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ E_i^{\mu}(t),\ \overline{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ \overline{E_i^{\mu}(t)}\right)$$



- We want to know the probability of the empirical mean deviating far from its true mean.

Proving the Regret Bound : Analyzing the Third Term

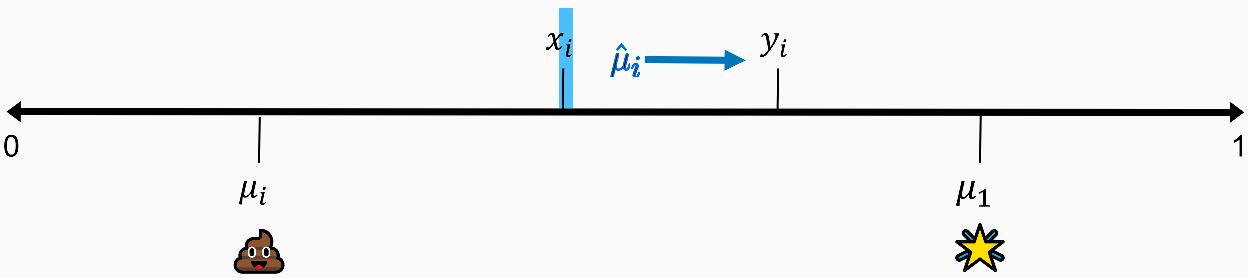
$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ E_i^{\mu}(t),\ E_i^{\theta}(t)\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ E_i^{\mu}(t),\ \overline{E_i^{\theta}(t)}\right) + \boxed{\sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ \overline{E_i^{\mu}(t)}\right)}$$



- We want to know the probability of the empirical mean deviating far from its true mean.
- Recall from last lecture Chernoff-Hoeffding bound provides an upper bound on this probability.

Proving the Regret Bound : Analyzing the Third Term

$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \overline{E_i^{\theta}(t)}\right) + \boxed{\sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{\overline{E_i^{\mu}(t)}}\right)}$$



- We want to know the probability of the empirical mean deviating far from its true mean.
- Recall from last lecture Chernoff-Hoeffding bound provides an upper bound on this probability.

$$\sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \overline{E_i^{\mu}(t)}\right) \leq \frac{1}{d(x_i, \mu_i)} + 1$$

Proving the Regret Bound : Putting Everything Together

$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ E_i^{\mu}(t),\ E_i^{\theta}(t)\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ E_i^{\mu}(t),\ \overline{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ \overline{E_i^{\mu}(t)}\right)$$

Proving the Regret Bound : Putting Everything Together

$$\mathbb{E}[N_i(T)] = \boxed{\sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)} \right)} + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \overline{E_i^{\theta}(t)} \right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \overline{E_i^{\mu}(t)} \right)$$

$$\mathbb{E}[N_i(T)] \leq \boxed{O(1)} + \ldots$$

Proving the Regret Bound : Putting Everything Together

$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ E_i^{\mu}(t),\ E_i^{\theta}(t)\right) + \boxed{\sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ E_i^{\mu}(t),\ \overline{E_i^{\theta}(t)}\right)} + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ \overline{E_i^{\mu}(t)}\right)$$

$$\mathbb{E}[N_i(T)] \leq O(1) + \boxed{\frac{\log T}{d(x_i, y_i)} + 1} + \dots$$

Proving the Regret Bound : Putting Everything Together

$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \; E_i^{\mu}(t), \; E_i^{\theta}(t)\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \; E_i^{\mu}(t), \overline{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \overline{E_i^{\mu}(t)}\right)$$

$$\mathbb{E}[N_i(T)] \leq O(1) + \frac{\log T}{d(x_i, y_i)} + 1 + \frac{1}{d(x_i, \mu_i)} + 1$$

Proving the Regret Bound : Putting Everything Together

$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ E_i^{\mu}(t),\ E_i^{\theta}(t)\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ E_i^{\mu}(t),\ \overline{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i,\ \overline{E_i^{\mu}(t)}\right)$$

$$\mathbb{E}[N_i(T)] \leq O(1) + \frac{\log T}{d(x_i, y_i)} + 1 + \frac{1}{d(x_i, \mu_i)} + 1$$

Proving the Regret Bound : Putting Everything Together

$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \overline{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \overline{E_i^{\mu}(t)}\right)$$

$$\mathbb{E}[N_i(T)] \leq O(1) + \frac{\log T}{d(x_i, y_i)} + 1 + \frac{1}{d(x_i, \mu_i)} + 1$$

- Time to set the values of x_i and y_i.

Proving the Regret Bound : Putting Everything Together

$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \overline{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t) = i, \overline{E_i^{\mu}(t)}\right)$$

$$\mathbb{E}[N_i(T)] \leq O(1) + \frac{\log T}{d(x_i, y_i)} + 1 + \frac{1}{d(x_i, \mu_i)} + 1$$

- Time to set the values of x_i and y_i.
- Set x_i and y_i such that for some $\epsilon = [0, 1]$, $d(x_i, \mu_1) = \frac{d(\mu_i, \mu_1)}{1+\epsilon}$ and $d(x_i, y_i) = \frac{d(\mu_i, \mu_1)}{(1+\epsilon)^2}$

Proving the Regret Bound : Putting Everything Together

$$\mathbb{E}[N_i(T)] = \sum_{t=1}^{T} \mathbb{P}\left(i(t)=i,\ E_i^{\mu}(t),\ E_i^{\theta}(t)\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t)=i,\ E_i^{\mu}(t),\ \overline{E_i^{\theta}(t)}\right) + \sum_{t=1}^{T} \mathbb{P}\left(i(t)=i,\ \overline{E_i^{\mu}(t)}\right)$$

$$\mathbb{E}[N_i(T)] \leq O(1) + \frac{\log T}{d(x_i, y_i)} + 1 + \frac{1}{d(x_i, \mu_i)} + 1$$

- Time to set the values of x_i and y_i.
- Set x_i and y_i such that for some $\epsilon = [0, 1]$, $d(x_i, \mu_1) = \frac{d(\mu_i, \mu_1)}{1+\epsilon}$ and $d(x_i, y_i) = \frac{d(\mu_i, \mu_1)}{(1+\epsilon)^2}$

$$\mathbb{E}[N_i(T)] \leq (1+\epsilon)^2 \frac{\log T}{d(\mu_i, \mu_1)} + O\left(\frac{K}{\epsilon^2}\right)$$

Proving the Regret Bound : Final Step

Expected cumulative regret after T time steps is

$$\begin{aligned}
\mathfrak{R}(T) &= \sum_i \Delta_i \, \mathbb{E}[N_i(T)] \\
&\le (1+\epsilon)^2 \sum_i \frac{\log T}{d(\mu_i, \mu_1)} \Delta_i + O\left(\frac{K}{\epsilon^2}\right) \quad \square
\end{aligned}$$

Distribution-free Regret Bound for Thompson Sampling

Theorem (Theorem 2 from Agrawal and Goyal [2013])

After T time steps, the expected cumulative regret of Thompson sampling using Beta priors is

$$Regret = \mathfrak{R}(T) \leq O(\sqrt{KT \log(T)})$$

Summary

- Solving Bandits using a Bayesian Perspective.

Summary

- Solving Bandits using a Bayesian Perspective.
- Thompson Sampling and its Regret Bound.

Summary

- Solving Bandits using a Bayesian Perspective.
- Thompson Sampling and its Regret Bound.
- Proof for the Regret Bound.

References

Shipra Agrawal and Navin Goyal. Further optimal regret bounds for thompson sampling. In *Proceedings of the Sixteenth International Conference on Artificial Intelligence and Statistics*, pages 99–107, 2013. URL http://proceedings.mlr.press/v31/agrawal13a.pdf.

Peter Auer, Nicolò Cesa-Bianchi, and Paul Fischer. Finite-time analysis of the multiarmed bandit problem. *Mach. Learn.*, 47(2–3):235–256, may 2002. ISSN 0885-6125. doi: 10.1023/A:1013689704352. URL https://doi.org/10.1023/A:1013689704352.

Extra Material

- For more insights into Thompson Sampling, watch this video (till minute 32).
- Some resources on frequentist and Bayesian perspective : Stanford Encyclopedia of Philosophy articles - Interpretations of Probability by Alan Hájek, and Philosophy of Statistics by Jan-Willem Romeijn, a StackExchange question.
- For the purpose of producing useful and self-consistent results, any frequentist interpretation can generally be given a Bayesian interpretation, and vice versa.

Next lecture

- Non-stationary Stochastic Bandits.
- Adversarial Bandits.
- Dueling Bandits (and a lower bound).
- Contextual Bandits.

Main Lemma

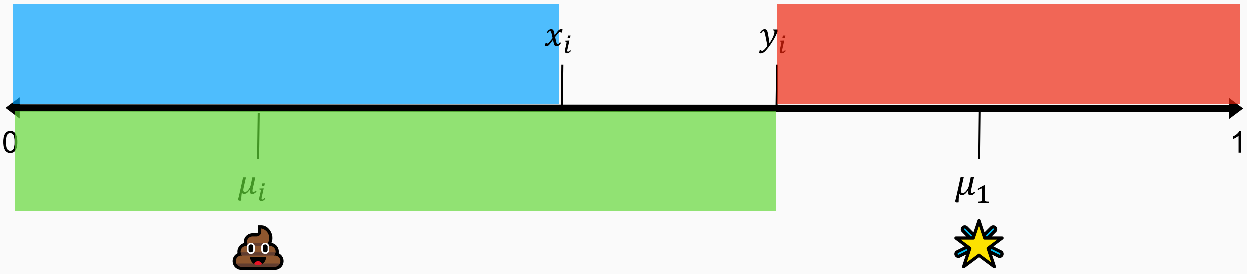
- Conditioned on any history, $\mathbb{P}$(playing any suboptimal arm at t) $\leq$ linear function of $\mathbb{P}$(playing the optimal arm at t).

Lemma (Lemma 1 from Agrawal and Goyal [2013])

For all $t = 1, \ldots, T$ and all suboptimal arms i i.e. $i \neq 1$,

$$\mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)} \mid \mathcal{F}_{t-1}\right)$$
$$\leq \frac{(1 - p_{i,t})}{p_{i,t}} \mathbb{P}\left(i(t) = 1, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)} \mid \mathcal{F}_{t-1}\right)$$

where $p_{i,t} := \mathbb{P}(\boxed{\theta_1(t) > y_i} \mid \mathcal{F}_{t-1})$



Proving the Main Lemma I

Lemma (Main Lemma)

For all $t = 1, \ldots, T$ and all suboptimal arms i i.e. $i \neq 1$,

$$\mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)} \mid \mathcal{F}_{t-1}\right)$$
$$\leq \frac{(1 - p_{i,t})}{p_{i,t}} \mathbb{P}\left(i(t) = 1, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)} \mid \mathcal{F}_{t-1}\right)$$

Proving the Main Lemma I

Lemma (Main Lemma)

For all $t = 1, \ldots, T$ and all suboptimal arms i i.e. $i \neq 1$,

$$\mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)} \mid \mathcal{F}_{t-1}\right) \leq \frac{(1 - p_{i,t})}{p_{i,t}} \mathbb{P}\left(i(t) = 1, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)} \mid \mathcal{F}_{t-1}\right)$$

Proof.

- History till time $t-1$ i.e. $\mathcal{F}_{t-1}$ determines the value of $\boxed{E_i^{\mu}(t)}$.

Proving the Main Lemma I

Lemma (Main Lemma)

For all $t = 1, \ldots, T$ and all suboptimal arms i i.e. $i \neq 1$,

$$\mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)} \mid \mathcal{F}_{t-1}\right)$$
$$\leq \frac{(1 - p_{i,t})}{p_{i,t}} \mathbb{P}\left(i(t) = 1, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)} \mid \mathcal{F}_{t-1}\right)$$

Proof.

- History till time $t-1$ i.e. $\mathcal{F}_{t-1}$ determines the value of $\boxed{E_i^{\mu}(t)}$.
- If $\mathcal{F}_{t-1}$ is such that $\boxed{E_i^{\mu}(t)}$ is false, then LHS is 0 and the lemma is trivially true as the RHS will also be 0.

Proving the Main Lemma I

Lemma (Main Lemma)

For all $t = 1, \ldots, T$ and all suboptimal arms i i.e. $i \neq 1$,

$$\mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)} \mid \mathcal{F}_{t-1}\right)$$
$$\leq \frac{(1 - p_{i,t})}{p_{i,t}} \mathbb{P}\left(i(t) = 1, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)} \mid \mathcal{F}_{t-1}\right)$$

Proof.

- History till time $t-1$ i.e. $\mathcal{F}_{t-1}$ determines the value of $\boxed{E_i^{\mu}(t)}$.
- If $\mathcal{F}_{t-1}$ is such that $\boxed{E_i^{\mu}(t)}$ is false, then LHS is 0 and the lemma is trivially true as the RHS will also be 0.
- So we try to prove the lemma when $\mathcal{F}_{t-1}$ is such that $\boxed{E_i^{\mu}(t)}$ is true,

Proving the Main Lemma I

Lemma (Main Lemma)

For all $t = 1, \ldots, T$ and all suboptimal arms i i.e. $i \neq 1$,

$$\mathbb{P}\left(i(t) = i, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)} \mid \mathcal{F}_{t-1}\right)$$
$$\leq \frac{(1 - p_{i,t})}{p_{i,t}} \mathbb{P}\left(i(t) = 1, \boxed{E_i^{\mu}(t)}, \boxed{E_i^{\theta}(t)} \mid \mathcal{F}_{t-1}\right)$$

Proof.

- History till time $t-1$ i.e. $\mathcal{F}_{t-1}$ determines the value of $\boxed{E_i^{\mu}(t)}$.
- If $\mathcal{F}_{t-1}$ is such that $\boxed{E_i^{\mu}(t)}$ is false, then LHS is 0 and the lemma is trivially true as the RHS will also be 0.
- So we try to prove the lemma when $\mathcal{F}_{t-1}$ is such that $\boxed{E_i^{\mu}(t)}$ is true, i.e. prove that $\mathbb{P}\left(i(t) = i \mid \boxed{E_i^{\theta}(t)}, \mathcal{F}_{t-1}\right) \leq \frac{(1-p_{i,t})}{p_{i,t}} \mathbb{P}\left(i(t) = 1 \mid \boxed{E_i^{\theta}(t)}, \mathcal{F}_{t-1}\right)$.

Proving the Main Lemma II

To prove: $\mathbb{P}\left(i(t) = i \mid \boxed{E_i^{\theta}(t)}, \mathcal{F}_{t-1}\right) \le \frac{(1-p_{i,t})}{p_{i,t}} \mathbb{P}\left(i(t) = 1 \mid \boxed{E_i^{\theta}(t)}, \mathcal{F}_{t-1}\right)$.

- $\boxed{E_i^{\theta}(t)}$ is $\theta_i(t) \le y_i$.

Proving the Main Lemma II

To prove: $\mathbb{P}\left(i(t) = i \mid E_i^\theta(t), \mathcal{F}_{t-1}\right) \leq \frac{(1-p_{i,t})}{p_{i,t}} \mathbb{P}\left(i(t) = 1 \mid E_i^\theta(t), \mathcal{F}_{t-1}\right)$.

- $E_i^\theta(t)$ is $\theta_i(t) \leq y_i$.
- Selection rule of TS : select arm $i(t) = \arg\max_j \theta_j(t)$.

Proving the Main Lemma II

To prove: $\mathbb{P}\left(i(t) = i \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right) \leq \frac{(1-p_{i,t})}{p_{i,t}} \mathbb{P}\left(i(t) = 1 \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$.

- $\boxed{E_i^\theta(t)}$ is $\theta_i(t) \leq y_i$.
- Selection rule of TS : select arm $i(t) = \arg\max_j \theta_j(t)$.
- So given $\boxed{E_i^\theta(t)}$, $i(t) = i$ only if $\theta_b(t) \leq y_i, \forall$ arms b

Proving the Main Lemma II

To prove: $\mathbb{P}\left(i(t)=i \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right) \leq \frac{(1-p_{i,t})}{p_{i,t}} \mathbb{P}\left(i(t)=1 \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$.

- $\boxed{E_i^\theta(t)}$ is $\theta_i(t) \leq y_i$.
- Selection rule of TS : select arm $i(t) = \arg\max_j \theta_j(t)$.
- So given $\boxed{E_i^\theta(t)}$, $i(t) = i$ only if $\theta_b(t) \leq y_i, \forall$ arms b

$\mathbb{P}\left(i(t)=i \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$

Proving the Main Lemma II

To prove: $\mathbb{P}\left(i(t) = i \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right) \leq \frac{(1-p_{i,t})}{p_{i,t}} \mathbb{P}\left(i(t) = 1 \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$.

- $\boxed{E_i^\theta(t)}$ is $\theta_i(t) \leq y_i$.
- Selection rule of TS : select arm $i(t) = \arg\max_j \theta_j(t)$.
- So given $\boxed{E_i^\theta(t)}$, $i(t) = i$ only if $\theta_b(t) \leq y_i, \forall$ arms b

$$\mathbb{P}\left(i(t) = i \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right) \leq \mathbb{P}\left(\theta_b(t) \leq y_i, \forall b \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$$

Proving the Main Lemma II

To prove: $\mathbb{P}\left(i(t) = i \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right) \le \frac{(1-p_{i,t})}{p_{i,t}} \mathbb{P}\left(i(t) = 1 \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$.

- $\boxed{E_i^\theta(t)}$ is $\theta_i(t) \le y_i$.
- Selection rule of TS : select arm $i(t) = \arg\max_j \theta_j(t)$.
- So given $\boxed{E_i^\theta(t)}$, $i(t) = i$ only if $\theta_b(t) \le y_i, \forall$ arms b

$$
\begin{aligned}
\mathbb{P}\left(i(t) = i \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right) &\le \mathbb{P}\left(\theta_b(t) \le y_i, \forall b \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right) \\
&= \mathbb{P}\left(\theta_1(t) \le y_i \mid \mathcal{F}_{t-1}\right) \\
&\quad \cdot \mathbb{P}\left(\theta_b(t) \le y_i, \forall b \ne 1 \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)
\end{aligned}
$$

Proving the Main Lemma II

To prove: $\mathbb{P}\left(i(t) = i \mid E_i^\theta(t), \mathcal{F}_{t-1}\right) \le \frac{(1-p_{i,t})}{p_{i,t}} \mathbb{P}\left(i(t) = 1 \mid E_i^\theta(t), \mathcal{F}_{t-1}\right)$.

- $E_i^\theta(t)$ is $\theta_i(t) \le y_i$.
- Selection rule of TS : select arm $i(t) = \arg\max_j \theta_j(t)$.
- So given $E_i^\theta(t)$, $i(t) = i$ only if $\theta_b(t) \le y_i, \forall$ arms b

$$
\begin{aligned}
\mathbb{P}\left(i(t) = i \mid E_i^\theta(t), \mathcal{F}_{t-1}\right) &\le \mathbb{P}\left(\theta_b(t) \le y_i, \forall b \mid E_i^\theta(t), \mathcal{F}_{t-1}\right) \\
&= \mathbb{P}\left(\theta_1(t) \le y_i \mid \mathcal{F}_{t-1}\right) \\
&\quad \cdot \mathbb{P}\left(\theta_b(t) \le y_i, \forall b \ne 1 \mid E_i^\theta(t), \mathcal{F}_{t-1}\right) \\
&= (1 - p_{i,t}) \cdot \mathbb{P}\left(\theta_b(t) \le y_i, \forall b \ne 1 \mid E_i^\theta(t), \mathcal{F}_{t-1}\right)
\end{aligned}
$$

Using $p_{i,t} \coloneqq \mathbb{P}(\theta_1(t) > y_i \mid \mathcal{F}_{t-1})$

Proving the Main Lemma III

$$\mathbb{P}\left(i(t) = i \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$$

$$\leq \frac{(1-p_{i,t})}{p_{i,t}} \cdot p_{i,t} \cdot \mathbb{P}\left(\theta_b(t) \leq y_i, \forall b \neq 1 \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$$

Proving the Main Lemma III

$$\mathbb{P}\left(i(t) = i \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$$
$$\leq \frac{(1 - p_{i,t})}{p_{i,t}} \cdot p_{i,t} \cdot \mathbb{P}\left(\theta_b(t) \leq y_i, \forall b \neq 1 \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$$

$$p_{i,t} \cdot \mathbb{P}\left(\theta_b(t) \leq y_i, \forall b \neq 1 \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$$

Proving the Main Lemma III

$$\mathbb{P}\left(i(t) = i \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$$
$$\leq \frac{(1 - p_{i,t})}{p_{i,t}} \cdot p_{i,t} \cdot \mathbb{P}\left(\theta_b(t) \leq y_i, \forall b \neq 1 \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$$

$$p_{i,t} \cdot \mathbb{P}\left(\theta_b(t) \leq y_i, \forall b \neq 1 \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$$
$$\leq \mathbb{P}(\theta_1(t) > y_i \mid \mathcal{F}_{t-1}) \cdot \mathbb{P}\left(\theta_b(t) \leq y_i, \forall b \neq 1 \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$$

Proving the Main Lemma III

$$\mathbb{P}\left(i(t) = i \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$$
$$\leq \frac{(1 - p_{i,t})}{p_{i,t}} \cdot p_{i,t} \cdot \mathbb{P}\left(\theta_b(t) \leq y_i, \forall b \neq 1 \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$$

$$p_{i,t} \cdot \mathbb{P}\left(\theta_b(t) \leq y_i, \forall b \neq 1 \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$$
$$\leq \mathbb{P}(\theta_1(t) > y_i \mid \mathcal{F}_{t-1}) \cdot \mathbb{P}\left(\theta_b(t) \leq y_i, \forall b \neq 1 \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$$
$$= \mathbb{P}\left(\theta_b(t) \leq y_i < \theta_1(t), \forall b \neq 1 \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$$

Proving the Main Lemma III

$$\mathbb{P}\left(i(t) = i \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$$
$$\leq \frac{(1 - p_{i,t})}{p_{i,t}} \cdot p_{i,t} \cdot \mathbb{P}\left(\theta_b(t) \leq y_i, \forall b \neq 1 \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$$

$$p_{i,t} \cdot \mathbb{P}\left(\theta_b(t) \leq y_i, \forall b \neq 1 \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$$
$$\leq \mathbb{P}(\theta_1(t) > y_i \mid \mathcal{F}_{t-1}) \cdot \mathbb{P}\left(\theta_b(t) \leq y_i, \forall b \neq 1 \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$$
$$= \mathbb{P}\left(\theta_b(t) \leq y_i < \theta_1(t), \forall b \neq 1 \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right)$$
$$\leq \mathbb{P}\left(i(t) = 1 \mid \boxed{E_i^\theta(t)}, \mathcal{F}_{t-1}\right) \quad \square$$

Appendix I

Assessment : Empty Assessment Form

Research Project Assessment Form

Department of Computer Science (TU Eindhoven)

Track & group		**Course code**	2AMM20
Project supervisor			
Title			

Assessment of the various aspects of the research project

		assessment				
		U	S	G	VG	E
Results	quality of the results					
	quantity of the results					
	complexity of the problem					
	Explanation:					
Report	structure					
	discussion of related work and context					
	clarity of presentation and correctness of arguments					
	English usage					
	general appearance (layout, figures and tables, etcetera)					
	Explanation:					
Execution	independence in execution of the project					
	independence in writing the report					
	planning and meeting deadlines					
	communication					
	Explanation:					

Overall assessment of the research project

explanation	group grade

Modifications to the group grade

explanation	modification

Final grades

name	grade

Explanation of Assessment Form

Research projects are judged on three criteria: results, report, and execution. For each of these criteria there are several subcriteria, which are scored on the following scale:

U = unsatisfactory
S = sufficient
G = good
VG = very good
E = excellent

Note: Students are expected to be able to do good work. The score "good" thus represents what can expected from a normal student; it does not imply above-average results.

The scores and the overall performance with respect to each of the criteria together determine the final grade for the graduation project. There is no fixed scheme for this, but the following serves as a guideline for arriving at the final grade. Note that grades need not be integers, halves are also allowed. Also note that a passing grade is 6.0 or higher.

Grade	Typical scores and evaluation
3	The work is unsatisfactory on aspects concerning results (in particular on quality or quantity of results) or report (in particular on structure, or clarity and correctness), or on many aspects overall.
6	The work scores satisfactory (and not more) on aspects concerning results and concerning the report, and typically also on independence in execution. The remaining scores are good at best.
7	There may be some scores that are only satisfactory, but most scores are good.
8	The work is very good with respect to several criteria and good with respect to the remaining ones. Typically, a solid piece of work with interesting although perhaps not very surprising results, achieved with a reasonable level of independence.
9	The work is excellent with respect to several criteria and very good with respect to the remaining criteria. The thesis presents an innovative solution to a complex problem, obtained with a high level of independence. For research-oriented projects, the work can lead to a publication in a good conference or journal; for design-oriented projects the work can be directly, or with relatively little effort, be applied in an industrial context and/or a concrete product (e.g., integrated into a large software system).
10	The work is excellent with respect to all five criteria and the work is clearly outstanding with respect to quantity or quality. For research-oriented projects, the work can lead to two publications in good conferences or one publication in a top conference or journal.

Appendix J

Assessment : Assessment Form with High Grade

Research Project Assessment Form

Department of Computer Science (TU Eindhoven)

Track & group	RL group 5	**Course code**	2AMM20
Project supervisor	dr. Pratik Gajane		
Title	Generalizing distribution of partial rewards for multi-armed bandits with temporally-partitioned rewards		

Assessment of the various aspects of the research project

		assessment				
		U	S	G	VG	E
Results	quality of the results				×	
	quantity of the results				×	
	complexity of the problem				×	

Explanation:

This is where the group really excelled. Their problem setting is a natural and well-justified extension to an existing problem. The proposed algorithm and the corresponding regret bound are a novel contribution to the literature.

		U	S	G	VG	E
Report	structure			×		
	discussion of related work and context				×	
	clarity of presentation and correctness of arguments				×	
	English usage				×	
	general appearance (layout, figures and tables, etcetera)				×	

Explanation:

The organization of the first two sections could be improved. Perhaps, they need to be shortened too. The theoretical background section has a couple of minor issues, e.g., use of \mu_i without defining it first,use of T_j(n) without defining it first, use of "machine" instead of "arm" without clarifying what a "machine" means in this context, imprecise characterization for UCB1 results (use of "random variable" instead of specifying "mean reward" instead, the rewards don't need to be in [0,1] as stated, they only need to be bounded). The explanation of alpha-smoothness above Def 2.1 could be made clearer. \bar{R}^i is used in Def 2.1 but it is defined later in Section 3. Some of the used notation is inconsistent, overloaded or undefined.

		U	S	G	VG	E
Execution	independence in execution of the project				×	
	independence in writing the report					×
	planning and meeting deadlines					×
	communication					×

Explanation:

They worked independently seeking some advice and feedback when they needed it. While seeking advice with the mathematical analysis, their questions were non-trivial and it seemed like they had made an attempt to solve the problems themselves before seeking my help. They implemented most of my feedback in time and the communication with them was unambiguous.

Overall assessment of the research project

explanation	group grade
With some improvement in the writing and a few experimental results (which were outside the purview of this course), this could very well be a paper at a good conference.	9

Modifications to the group grade

explanation	modification
Main deadline bonus Peer review : consensus that all group members contributed equally.	+1 for all

Final grades

name	grade
[redacted]	10 for all

Explanation of Assessment Form

Research projects are judged on three criteria: results, report, and execution. For each of these criteria there are several subcriteria, which are scored on the following scale:

U = unsatisfactory
S = sufficient
G = good
VG = very good
E = excellent

Note: Students are expected to be able to do good work. The score "good" thus represents what can expected from a normal student; it does not imply above-average results.

The scores and the overall performance with respect to each of the criteria together determine the final grade for the graduation project. There is no fixed scheme for this, but the following serves as a guideline for arriving at the final grade. Note that grades need not be integers, halves are also allowed. Also note that a passing grade is 6.0 or higher.

Grade	Typical scores and evaluation
3	The work is unsatisfactory on aspects concerning results (in particular on quality or quantity of results) or report (in particular on structure, or clarity and correctness), or on many aspects overall.
6	The work scores satisfactory (and not more) on aspects concerning results and concerning the report, and typically also on independence in execution. The remaining scores are good at best.
7	There may be some scores that are only satisfactory, but most scores are good.
8	The work is very good with respect to several criteria and good with respect to the remaining ones. Typically, a solid piece of work with interesting although perhaps not very surprising results, achieved with a reasonable level of independence.
9	The work is excellent with respect to several criteria and very good with respect to the remaining criteria. The thesis presents an innovative solution to a complex problem, obtained with a high level of independence. For research-oriented projects, the work can lead to a publication in a good conference or journal; for design-oriented projects the work can be directly, or with relatively little effort, be applied in an industrial context and/or a concrete product (e.g., integrated into a large software system).
10	The work is excellent with respect to all five criteria and the work is clearly outstanding with respect to quantity or quality. For research-oriented projects, the work can lead to two publications in good conferences or one publication in a top conference or journal.

Appendix K

Assessment : Assessment Form with Low Grade

Research Project Assessment Form

Department of Computer Science (TU Eindhoven)

Track & group	RL group 02	Course code	2AMM20
Project supervisor	dr. Pratik Gajane		
Title	Modifying Thompson Sampling for Sleeping Bandits with Switching Costs		

Assessment of the various aspects of the research project

		assessment				
		U	S	G	VG	E
Results	quality of the results	X				
	quantity of the results	X				
	complexity of the problem		X			
	Explanation: The proposed algorithm does not make use of the switching costs which is the key feature of the considered problem. Hence, the resulting bounds too are not in terms of the switching costs. The proven results and the mathematical analysis are both quite simplistic. Since, advanced mathematical analysis was extensively taught in the course, it was expected that the mathematical analysis in the research projects would be sophisticated too. However, that is not the case with this project.					
Report	structure		X			
	discussion of related work and context	X				
	clarity of presentation and correctness of arguments		X			
	English usage		X			
	general appearance (layout, figures and tables, etcetera)		X			
	Explanation: Related work section is missing. There are only 5 references out of which only one is briefly discussed. The imprecise mathematical analysis at times display a lack of understanding of the topic. Everything else is sufficient and no more.					
Execution	independence in execution of the project		X			
	independence in writing the report		X			
	planning and meeting deadlines		X			
	communication		X			
	Explanation: The group showed inadequate progress in the initial few weeks and took too long to formulate their problem. The progress reports sent by the group were not detailed (and once they did not even send a progress report). The discussions during the group meetings were mostly unproductive.					

Overall assessment of the research project

explanation	group grade
The results, the report and the execution were bare minimum. I do appreciate that they pulled through with a report in the end, and I am taking into account here that they effectively were with only four people to a group.	5.5

Modifications to the group grade

explanation	modification
Peer review from four group members mention that the other group member did not contribute at all. So the student who did not contribute at all receives an incomplete grade for the course, while others receive +0.5.	

Final grades

name	grade
[redacted]	6
[redacted]	6
[redacted]	6
[redacted]	6
[redacted]	Incomplete

Explanation of Assessment Form

Research projects are judged on three criteria: results, report, and execution. For each of these criteria there are several subcriteria, which are scored on the following scale:

U = unsatisfactory
S = sufficient
G = good
VG = very good
E = excellent

Note: Students are expected to be able to do good work. The score "good" thus represents what can expected from a normal student; it does not imply above-average results.

The scores and the overall performance with respect to each of the criteria together determine the final grade for the graduation project. There is no fixed scheme for this, but the following serves as a guideline for arriving at the final grade. Note that grades need not be integers, halves are also allowed. Also note that a passing grade is 6.0 or higher.

Grade	Typical scores and evaluation
3	The work is unsatisfactory on aspects concerning results (in particular on quality or quantity of results) or report (in particular on structure, or clarity and correctness), or on many aspects overall.
6	The work scores satisfactory (and not more) on aspects concerning results and concerning the report, and typically also on independence in execution. The remaining scores are good at best.
7	There may be some scores that are only satisfactory, but most scores are good.
8	The work is very good with respect to several criteria and good with respect to the remaining ones. Typically, a solid piece of work with interesting although perhaps not very surprising results, achieved with a reasonable level of independence.
9	The work is excellent with respect to several criteria and very good with respect to the remaining criteria. The thesis presents an innovative solution to a complex problem, obtained with a high level of independence. For research-oriented projects, the work can lead to a publication in a good conference or journal; for design-oriented projects the work can be directly, or with relatively little effort, be applied in an industrial context and/or a concrete product (e.g., integrated into a large software system).
10	The work is excellent with respect to all five criteria and the work is clearly outstanding with respect to quantity or quality. For research-oriented projects, the work can lead to two publications in good conferences or one publication in a top conference or journal.

Appendix L

Assessment : How Peer Review Works

Overall assessment of the research project

explanation	group grade
Even if the proven results are barely sufficient, they displayed enough competence in other aspects of the project to pass the course. . .	6.5

Modifications to the group grade

explanation	modification
Main deadline bonus Peer review : The individual peer reviews suggest the following levels of contribution: [redacted] > [redacted] = [redacted] = [redacted] > [redacted] So Jelle receives +0.5 and Georgi and Niek receives -0.5.	+1 for all

Final grades

name	grade
[redacted]	7.5
[redacted]	8
[redacted]	7
[redacted]	7.5
[redacted]	7.5

Appendix M

Support Letter from Dr. Wouter Duivesteijn

Technische Universiteit
Eindhoven
University of Technology

Data Mining Group

Navigation address: De Zaale, Eindhoven
P.O. Box 513, 5600 MB Eindhoven
the Netherlands
Internal address: 5 MetaForum 7.145
www.tue.nl

5 MetaForum 7.145, P.O. Box 513, 5600 MB Eindhoven, NL
To whom it may concern

Date
September 13, 2023

Contact
dr. W. Duivesteijn
T +31 40 247 4008
w.duivesteijn@tue.nl

To whom it may concern,

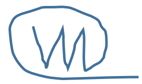
In the previous academic year, 2022/2023, dr. Pratik Gajane joined me as a co-lecturer in my course 2AMM20: Research Topics in Data Mining. The concept of the course is that we divide a large cohort of new (first year, first block) MSc students into three tracks, and within these tracks into groups of five; each group's final project is to contribute to the body of scientific knowledge, by writing a research paper. They must identify a knowledge gap, bridge it, and write up their results in a scientific paper that would be in a form submittable to an international data mining conference. The tasks of dr. Gajane were to create a track on a topic of his choice, deliver lectures in weeks 1-3 of the block that teach the students in his track all the basics required to get up to speed with the scientific state-of-the-art in the topic of the track, and subsequently (in weeks 4-8) supervising the groups in his track on their research project trajectory.

I have been teaching this course in this form now for two editions (the third is ongoing), and so far we have supervised 61 groups across tracks on five distinct themes. Dr. Gajane developed a track on the topic of Reinforcement Learning, and completely independently created the required five lectures covering the content. He asked for my feedback in the creation of a few of these lectures, but it became rather immediately apparent to me that my feedback would only be useful in finetuning the lectures: dr. Gajane created lecture material that was already of a very high standard without outside help. It is a pleasure to have a co-lecturer that can perform at a high level independently. The students agreed with me on this matter: the official student evaluations included the following testimonials:
"The lessons of the first weeks were clear and interesting (for reinforcement learning). Pratik was nice and clear when explaining."
and:
"The topic seems difficult, but I think the teacher did an amazing job with the way all the information is organized on the slides (colors, variables, mathematical formulations, and equations), and all the real-life examples made the topic interesting. I think the teacher explained the content in a clear and comprehensive way and made it seem easier than it actually is. This course increased my interest in Reinforcement Learning. And if there is ever a course on this topic, I think Pratik would be great for it."

In this research-oriented course, the ideal outcome is if we can support the students during their research project so well that their results are good enough to lead to an actual joint scientific publication. Of the 61 groups we have seen in the course, 7 have delivered publishable results, and 4 have indeed now led to a published paper. One of those was from a group supervised by Pratik, out of the seven groups in his track. This speaks very well indeed of Pratik's teaching skills in terms of supervising small groups; his groups are overrepresented in my course's publication records.

Dr. Gajane would be always welcome to coordinate a track in my course: his teaching can be depended upon, students appreciate what he does, and his instincts regarding how teaching is supposed to be done are spot-on. His obligations prevented him from participating in my course in this academic year, but that is the only preventing factor: I would welcome a continuation of my collaboration with dr. Gajane with open arms.

Yours faithfully,
dr. W. Duivesteijn



Assistant Professor Data Mining at TU/e

Appendix N

Support Letter from Dr. Emilia Barakova

EINDHOVEN UNIVERSITY OF TECHNOLOGY

To Whom It May Concern

Dr. Ir. Emilia I. Barakova,
Faculty of Industrial Design,
Head TU/e Social Robotics Lab
Eindhoven University of Technology
Email: e.i.barakova@tue.nl

Evaluation of the teaching performance of Dr. Pratik Gajane

I am happy to provide a positive evaluation of Dr. Pratik Gajane, who has been excellent support of the teaching team for the course "Embodying intelligent behavior on social context" during the Q1, 2022-2023.

Since Pratik joined our teaching team, I have had the pleasure of closely observing his work. I am pleased to report that he has consistently met or exceeded our expectations in several key areas, such as teaching one lecture and giving two practicums and being an invaluable support to the students when they were preparing their deliverables (a small application that required a proper choosing and implementing of a learning algorithm in societal context). The student needed support with conceptual thinking and programming challenges, which Pratik provided in a great way.. The students explicitly mentioned his contribution in the course feedback. Pratik showed a deep understanding of the subject matter and could explain complex concepts in a clear manner which has greatly enhanced the learning experience for our students.

Pratik has proven to be highly reliable and responsible. He was present at every lecture and meeting and has taken the initiative to identify and address student needs. His communication skills are good, still when giving lectures he may become nervous and his accent is becoming a bit difficult to follow. Also, he speaks a bit too fast which are points for improvement.

Pratik has been an integral part of our teaching team, working collaboratively with me to enhance the overall quality of the course. Their positive attitude and willingness to collaborate have been greatly appreciated. Even after the course finished, he was involved in helping a student team which performed best to write a short paper that was accepted at an international conference and the students had the experience of this process.

Based on the very good performance of Pratik and his contributions to the success of the course I can certainly give positive advice to the BKO committee to grant him the teaching certificate.

With Kind Regards,

Emilia I. Barakova 10.09.2023

Appendix O

Self Evaluation Form Filled on October 27, 2022

Feel free to leave certain criteria blank if you are unsure about your own experience and/or want to discuss it during the intake.

Competence 1: (re)designing teaching	1	2	3	4	5
1. Explain how your course is embedded in the curriculum or degree program as a whole	✓				
2. Design teaching based on the principles of 'constructive alignment'		✓			
3. Design active, effective, and efficient learning methods and learning materials			✓		
4. Design your teaching with respect to the specific (curricular) characteristics and needs of the students			✓		
5. Design your teaching in a practically and logistically feasible (do-able) way			✓		
Competence 2: teaching and supervision	1	2	3	4	5
1. Prepare an educational meeting: design and justify a lesson plan, create materials for a lesson activity			✓		
2. Conduct an educational meeting and reflect on your performance			✓		
3. Supervise students, individually and in groups			✓		
Competence 3: assessment	1	2	3	4	5
1. Design and implement the assessment of student development and learning outcomes		✓			
2. Analyze the assessment results and draw conclusions about validity, reliability, fairness and transparency			✓		
Competence 4: evaluating teaching	1	2	3	4	5
1. Conduct an evaluation and collect information (data) purposefully to improve your teaching		✓			
2. Analyze evaluation results, draw conclusions, and pinpoint areas for improvement		✓			
Competence 5: professionalization	1	2	3	4	5
1. Formulate your own vision on teaching and student learning			✓		
2. Manage your work as a teacher and can collaborate in a teaching team			✓		
3. Reflect on your work as a teacher and on your future professional development in teaching		✓			

During the UTQ program, you will develop a dossier where you demonstrate your developed abilities on all five teaching competences by providing products and validations and subsequently reflecting on what you did and the feedback you received. For more details about sub criteria, products and validations for each competence, please consult the TU/e UTQ handbook (chapter 4) – although this is not necessary prior to your UTQ intake.

Is there anything left you'd like to add regarding your educational experience, your teaching tasks, your self-assessment on the UTQ competences, your UTQ intake, or any other matters relating to educational professionalization?

Appendix P

Self Evaluation Form Filled on September 4, 2023

Feel free to leave certain criteria blank if you are unsure about your own experience and/or want to discuss it during the intake.

Competence 1: (re)designing teaching	1	2	3	4	5
1. Explain how your course is embedded in the curriculum or degree program as a whole			✓		
2. Design teaching based on the principles of 'constructive alignment'				✓	
3. Design active, effective, and efficient learning methods and learning materials				✓	
4. Design your teaching with respect to the specific (curricular) characteristics and needs of the students				✓	
5. Design your teaching in a practically and logistically feasible (do-able) way				✓	
Competence 2: teaching and supervision	1	2	3	4	5
1. Prepare an educational meeting: design and justify a lesson plan, create materials for a lesson activity					✓
2. Conduct an educational meeting and reflect on your performance				✓	
3. Supervise students, individually and in groups				✓	
Competence 3: assessment	1	2	3	4	5
1. Design and implement the assessment of student development and learning outcomes				✓	
2. Analyze the assessment results and draw conclusions about validity, reliability, fairness and transparency			✓		
Competence 4: evaluating teaching	1	2	3	4	5
1. Conduct an evaluation and collect information (data) purposefully to improve your teaching				✓	
2. Analyze evaluation results, draw conclusions, and pinpoint areas for improvement			✓		
Competence 5: professionalization	1	2	3	4	5
1. Formulate your own vision on teaching and student learning				✓	
2. Manage your work as a teacher and can collaborate in a teaching team				✓	
3. Reflect on your work as a teacher and on your future professional development in teaching			✓		

During the UTQ program, you will develop a dossier where you demonstrate your developed abilities on all five teaching competences by providing products and validations and subsequently reflecting on what you did and the feedback you received. For more details about sub criteria, products and validations for each competence, please consult the TU/e UTQ handbook (chapter 4) – although this is not necessary prior to your UTQ intake.

Is there anything left you'd like to add regarding your educational experience, your teaching tasks, your self-assessment on the UTQ competences, your UTQ intake, or any other matters relating to educational professionalization?