Bachelor thesis

Artificial Intelligence

Radboud University

The effects of generalist and specialist scientists on scientific discovery

Author:
M.L.C. van der Meiden
s1018091
`mieke.vandermeiden@ru.nl`

First supervisor:
dr. M. Blokpoel
Donders Institute for Brain, Cognition and Behaviour
`m.blokpoel@donders.ru.nl`

Second supervisor:
prof. dr. I.J.E.I. van Rooij
Donders Institute for Brain, Cognition and Behaviour
`i.vanrooij@donders.ru.nl`

IN·DEI·NOMINE·FELICITER

January 25, 2022

Abstract

This paper will research the effects of different distributions of scientists on scientific discovery. Specifically, two scientist types will be researched: a generalist and a specialist. To do so, agent-based modelling will be used where the scientist types will be modelled according to previously done observational research. The agents will try to build their own version of an observed cognitive system, which is represented by a finite state transducer. They will keep creating new theories until one performs well enough to satisfy them. To see the effect each scientist type has their prominence in the population will be varied and both the speed of discovery and their theory size will be measured. These results will be plotted and interpreted within the scope of the model. Afterwards, these conclusions will be placed in a bigger scope where allowed and finally the research question will be looked at one final time.

Contents

1 Introduction

When pursuing a career in science one often works from a general bachelor's degree to a more specialised field through a master's and a PhD, to eventually end up working in one of the many specializations that are out there — whether that be in research or at a company. Way before this though, when science was only just being born, a scientist was much more than just a scientist. Often, they were also philosophers, artists, athletes, and much more. This traditional view of a scientist is rooted in the image of the *Renaissance man* [sic]; a person who seeks to gain knowledge about many subjects. This extreme contrast between your average scientist then and now, can still be visible to a lesser extent today between current day specialist and generalist scientists. The current day generalist obviously differs from their renaissance counterpart as the entire scientific field has changed, but there are still researchers who generalize rather than specialize even though it may not be the most straightforward path.

So, the generalist of today may not be the majority, but what effect do they have? And what would happen if more generalist than specialist scientists came into the current scientific world? The effect a different distribution of these scientists may have on both the quality and the speed of scientific discovery will be the main focus of this paper.

Research into the increased specialization of science has been done before and most often analyses the progression of science and scientific education in an observational way. The paper by Bode et al., 1949 describes how scientific generalists can help manage complicated problems and shows how these generalists can be educated using the current academic course structure. A newer paper by Wray, 2005 re-examines the causes of scientific specialization and notes the importance of communication between specializations to come to reach scientific discovery. Wray does so by looking at the discovery of the bacterial theory of ulcers, which required knowledge from two specialized fields. The knowledge had been around for a while, but collaboration, and above all, communication was needed to reach a conclusion.

Both these papers draw their conclusions by empirically studying the history of science and looking at individual occurrences. This is a more traditional approach, and can give many insights, but one is bounded by what has happened and how well everything was documented.

Taking a different approach to understanding science can be done by using computer models. These models then represent some part of the scientific world and by running simulations under different conditions the effects within the model can be found. These findings may then carefully be used to speculate about the real world. When using agents within these models, you speak of agent-based modelling (ABM) and although this is not a new technique it hasn't often been used to model science (Payette, 2012). A setup very similar to the one that will be used in this thesis was used by Rich et al., 2021 who took a look at the intractability of coming up with a reasonable explanation for a cognitive system. They proved that even with perfect, errorless and noiseless, data this problem is intractable, and from this it followed that there can not be one single approach to scientific discovery. Their setup consisted of a scientist (the agent) who observes a cognitive system. All the agent can observe are situations and the behaviours resulting from those situations. With that, their goal is to explain the inner workings of the cognitive system. This method of ABM where an agent tries to find an explanation of a cognitive system is what I will use to model the generalist and specialist scientists and explore two of many different approaches to scientific discovery.

Other relevant research was done by Devezer et al., 2019 who, similar to this paper, took a look at the distribution of different types of scientists. With their four types they looked at different approaches to replication studies and their effect on how fast the 'true' answer to a research question was converged upon. Their model however was not an agent-based one but a statistical one.

This paper will also take an ABM approach to try to answer the research question. There will be two types of agents: one representing a generalist scientist and one representing a specialist scientist. The agents will observe a cognitive system which will be modelled by a finite state transducer (FST), which is a specific type of automaton that transforms the input into output. This corresponds with the conceptual level of situations and resulting behaviours from a cognitive system. The agents' goal is to come up with an explanation of how the system works. This explanation is just another FST and is what the agent thinks the observed system looks like. Each agent will continue creating new explanations until one of their theorized systems performs well enough and then a solution has been found. By varying the amount of generalists and specialists in the entire population of agents, their effect on the population as a whole can be observed.

To fully explain what is going on, the next section will go over the agent setup in more detail. After this, four formal models will be discussed: that of the cognitive system, that of a base agent which will serve as a proof of concept, and that of a generalist and specialist scientist. These last two will be derived from the base agent and will be adjusted to fit their real world counterparts. Finally the overall environment will be defined in which both the cognitive system and the agents reside.

Next, these formal models will be translated to computer code to prepare for the eventual simulations. This part will cover how the agents work in practice and certain design choices will be explained. It is this section that really gets into the practical 'how' of the agents.

With all this groundwork done, the simulations can be setup and started. This chapter starts by doing some preliminary testing with the implementation to determine a reasonable setup for the actual simulations. Once these tests are done the general setup is discussed once more, including specific parameter values, and the testing can begin. For these simulations three main aspects of the agents and their theories are tracked: the satisfaction speed, the accuracy variety, and the theory variety. To finish off this section, some further testing will be done looking at the agents' performance against chance.

These results will then be summarized and interpreted further in the next section to finally draw conclusions about this research and look ahead at the future.

2 Formal models

A cognitive scientist's goal is to figure out how a certain cognitive system works. By observing different situations and people's behaviours on those situations they can construct their own theory of how the system works. New observations will now either confirm or contradict the scientist's theory according to whether it predicts the correct behaviours for each situation. When a theory fails to correctly do so, the scientist adjusts it and continues observing. Once a theory has been found that works reasonably well, it may take a very long time before the scientists observers a situation or behaviour that contradicts it. At this point the scientist may be contented and stop trying to find a suitable explanation for the system.
A scientist can never know whether their theory actually matches the system; even if the exact same situation-behaviour pattern is created by both system and theory, their inner structure might still be completely different. This is something the scientist knows and has to live with.

The above description covers a situation where there is only one scientist observing a system and is illustrated in figure 2.1. In reality there are often multiple scientists observing and studying the same problem. This, slightly more complex, multi-agent setup will be discussed in more detail later, but first the focus will be on individual parts.

To simulate these scientists observing situations and behaviours of a cognitive system, they have to be formalized. They will be modelled as agents which can later be implemented using computer code. Before being able to implement anything, it is vital that there is a good understanding of what it is that needs to be implemented. Namely, some agents living in an environment together with the cognitive system they want to understand. For the most part language is kept consistent and conceptual. This means 'scientist', '(observed) system', and 'theory' will be used the most but sometimes more technical terms are preferred. Therefore the following words will hereby be declared interchangeable: 'scientist' = 'agent' and 'system' = 'FST' = 'theory'.

Firstly, the cognitive system will be specified, explaining how FSTs work and how they will be used in the paper. After this, a base agent will be described, which functions as groundwork for the generalist and specialist agents which will be analysed in the section after. To put it all together, the chapter will finish with the environment specification showing how the agents and the cognitive system interact.

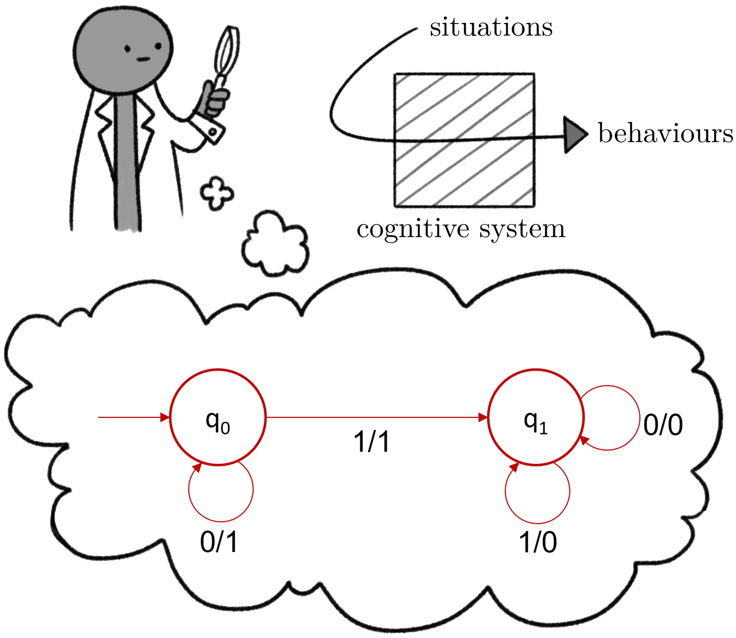


Figure 2.1: The situation where one scientists observes a cognitive system. They think of a representation of the inner workings of this system. In this case an FST.

2.1 System specification

Before giving the mathematical definition of an FST that will be used, a more conceptual explanation will be given. A transducer is a machine that has two tapes of infinite length: one from which it will read the input and one on which it will print the output. It can read and write one symbol at a time and by moving the tapes it can go through them one by one. The machine will read a symbol and together with its current state it will determine the output symbol and the next state. It will print this output on the output tape and then move both tapes to read the next input and create an empty spot for the next output. It will keep doing so, until it encounters an empty slot on the input tape. At that point the output tape has been filled and the input has been 'transduced'. Writing this down formally results in the following definition:

Definition 2.1.1. A system $M = \langle \Sigma, \Gamma, Q, q\_\_, F, \Delta \rangle$ is composed of an *input alphabet* Σ, an *output alphabet* Γ, a finite set of *states* Q, an *initial state* $q\_\_$, a set of *final states* F, and a set of *transitions* Δ satisfying

$$\Delta \subset Q \times \Sigma^* \times \Gamma^* \times Q$$

This definition has been taken from Berstel, 1979 with some symbols changed for this paper.

FSTs can be represented as graphs, in a very similar way to finite automata. Figure 2.2 shows the FST with the following definition: $\Sigma = \Gamma = \{0, 1\}, Q = \{q_0, q_1\}, q\_\_ = q_0, F = \{q_1\}$, and $\Delta = \{(q_0, 0, 1, q_0), (q_0, 1, 1, q_1), (q_1, 0, 0, q_1), (q_1, 1, 0, q_1)\}$. The initial state is marked by an ingoing arrow with no source and the final state is marked using a double circle. The input and output is written on the edges such that 'σ/γ' where $\sigma \in \Sigma$ and $\gamma \in \Gamma$.

Before moving on to the agent specification some further terms will be defined also, to avoid any confusion later on:

- **ingoing transition**: for a state q, an ingoing transition is a transition that goes into state q from some other state q'. Or more formally: (q', σ, γ, q).
- **outgoing transition**: for a state q, an outgoing transition is a transition that goes from q to some other state q'. Or more formally: (q, σ, γ, q'). The amount of

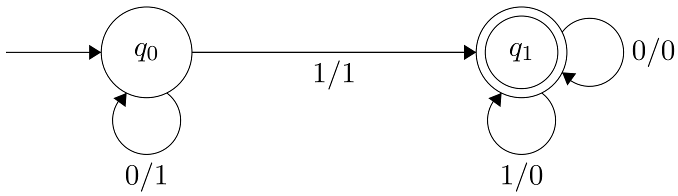


Figure 2.2: An example system represented by an finite state transducer.

outgoing transitions a state can have is limited by the amount of characters in the input alphabet. This is a result of using deterministic FSTs.

The specification of the agents will be in two main parts starting with the specification of the base agent. This section is vital for the further understanding of the other agents even though the base agent won't be used during simulations. Each of the agents will first be described conceptually, after which the formal model will be given.

2.2 Base scientist

The base scientist's behaviour is basic and random. They will look at a previously created theory and add onto it. How much their new addition is engrained within the old system is completely random, but it must have some connection. They will then test this new theory by looking at what situations result in the same behaviour as the observed system. Based on how well their theory works, the scientist will either continue adding onto previous theories or stop if they are contented. If they stop searching, a scientist will also tell the other scientists that they have found a solution they are satisfied with. Since every base scientist is the same, other scientists will also be satisfied and halt their searching. The found theory is presented and all's well that ends well.

Formally, an agent $a \in A$ will theorise a system M' which they hope to be equal to an observed system M. To do so, they have a data set $D \subset S \times B$ consisting of the data they observe from M. Each data point is a pair (s, b) of a situation $s \in S$ and a behaviour $b \in B$ where S and B are the sets of all possible situations and behaviours respectively. A theory M' can be used together with S to create behaviours B', which are the behaviours that would be if S occurred under M'. This will create a new dataset $D' \subseteq S \times B'$. By comparing the two datasets, the agent can determine how well their theory works. This comparison will yield a number c with $0 \leq c \leq 1$ where 0 indicates that no single situation generated the same behaviour and a 1 indicates that $D = D'$. Once c is high enough the agent will be satisfied and stop creating new theories. The agent has now created a *satisfactory theory*. How high c needs to be is determined by an acceptance threshold θ. Below is the mathematical definition for c:

Definition 2.2.1. The accuracy $c = \dfrac{\#\text{ of correct predictions}}{\#\text{ of predictions}}$ where a correct prediction is given by two pairs $(s, b) \in D$ and $(s', b') \in D'$ where $s = s'$ and $b = b'$.

The agent can make M' by taking a theory M'' created by another agent and add one state q to the FST structure. Note that M'' may consist of zero states. To connect this new state to the rest of the FST, the agent must add at least one ingoing transition. The definition for M' now becomes:

$$M' = \langle \Sigma, \Gamma, Q \cup q, q\_, F, \Delta \cup \Delta' \rangle$$

where

$$M'' = \langle \Sigma, \Gamma, Q, q\_, F, \Delta \rangle$$

M' has the same input and output alphabet and the same starting state. The newly added state gets added to the set of states Q and a new set of transitions Δ' from and to this state get added to the previous transitions. With this, the base agent has been defined.

2.3 Generalists and specialists

The generalists and specialists will be based upon the base scientist and retain quite a bit of its random behaviour. However, before going into the formal specification it is necessary to take a look at the conceptual differences between these two scientists. Looking at what defines these two scientists and where they differ in their approaches to science will provide a clear structure to build their more formal models. These conceptual differences will be discussed in the section below and will be based in literature. From this exploration, a choice will be made about what to model for each of the scientists and their formal definitions will be created.

2.3.1 Conceptual differences

When looking at types of scientists the definitions aren't binary. A specialist and generalist are opposites on a spectrum and most scientists will probably be somewhere in between. However, modelling a gradient with infinite variations is very difficult, if not impossible, and therefore abstracting to a binary system with just extreme generalists and specialists makes the modelling process a lot easier. Because of this abstraction, I will also be trying to find the largest, most defining aspects of each scientist type. To begin with a simple distinction: a generalist is someone with knowledge in multiple areas whereas a specialist focusses on just one domain. The total amount of knowledge they have would be the same, but a generalist has to spread this out over multiple areas, resulting in more surface-level knowledge, whereas a specialist can focus fully on one area, obtaining deep and thorough knowledge. They each choose to spend their time on a different aspect of scientific learning and one simply does not have time to do both.

Whether or not a scientist becomes a specialist is based on multiple factors such as education, scientific climate, but above all personal choice. Regardless of what the current situation calls for or what is easiest to do, one can always find way to adjust their tactics and change the 'type' of scientist they are (Bateman & Hess, 2015). This study by Bateman and Hess highlights the difference between scientists based on their motivation within science. A specialist scientist comes about when exploitation of current knowledge is more favourable than exploration of new fields. They are performance goal oriented and want to achieve results with little risks by staying within their current knowledge field and deepening it. A generalist is in favour of exploration and learning new things, and care less for the possible risks or low pay-off. This view is shared in the research by Teodoridis et al., 2019 which explored the different fields in which generalists

Generalist	**Specialist**
Surface-level knowledge of multiple fields	deep knowledge of a single field
exploration of new fields	exploitation of current field
novel ideas in stagnant fields	anticipation of change in fast-paced fields
mediator between different fields	theory builder

Table 2.1: A table summarizing the differences between a generalist and specialist scientist

and specialists would be more efficient. A stagnant field would benefit from a generalist, as their yearning for exploration and disregard for risk would help spark innovation by taking a look at different fields. This could possibly yield interesting novel ideas. On the contrary, a specialist would be suited to a fast-paced field as they have to be up to date and react to new changes; something that can only be achieved if your knowledge of the field is deep and thorough. This position of the generalist as exploring different fields and seeing the bigger picture may explain why they are sometimes referred to as philosophers of science who can see the totality obscured by specialization (Rice et al., 1950).

To summarize, on the surface a generalist has broad knowledge whereas a specialist has deep knowledge, but their practices also come down to a difference in scientific motivation and with that, a different approach. The differences discussed above have been summarized in table 2.1This difference in motivation is what will be modelled in the scientist agents also. A generalist will be able to take two theories and combine them, representing their ability to look at multiple different approaches. A specialist will work similarly to the base agent and build theories incrementally to represent their tactic of creating theories and continuously deepening their theories. The difference highlighted by Teodoris is an interesting one, but implementing a concept such as anticipation is not a task I will tackle in this paper.

The following two sections will cover the formal models of the generalist and specialist scientist. Here, their definitions in natural language will be given in a bit more detail, followed a more mathematical definition.

2.3.2 Generalist specification

The generalist scientist can combine two previously created theories. They will randomly select bits of information from the two theories and in doing so, create a new one. During this combination process some information may get lost. A generalist can take any theory they want as their widespread knowledge allows them to look across the entire field of science.

A generalist will theorise a system M' which they hope to be equal to an observed system M. Just like the base agent, they have access to a data set D which they can

use to test their theory.
The generalist agent will make M' by taking two previous theories M_1 and M_2 and combine them. They will do this by looking at each state pair with the same name and randomly select one state to keep in their new theory. The selected state, together with its outgoing transitions will make up the new system M'. When the sizes of the two combining systems don't match not all states will pair up. When this occurs a generalist will, for each of these lone states, select whether to keep it or not. M' now becomes:

$$M' = \langle \Sigma, \Gamma, Q' \subseteq Q_1 \uplus Q_2, q_{1\_} \wedge q_{2\_}, F' \subseteq Q', \Delta' \subseteq \Delta_1 \uplus \Delta_2 \rangle$$

where

$$M_1 = \langle \Sigma, \Gamma, Q_1, q_{1\_}, F_1, \Delta_1 \rangle \text{ and } M_2 = \langle \Sigma, \Gamma, Q_2, q_{2\_}, F_2, \Delta_2 \rangle$$

For M', the input and output alphabet stay the same. The set of states Q' is a subset of all the states of M_1 and M_2 and the same holds for the transition functions Δ'. Since the union of state and transition sets needs to contain duplicate states multisets have been used. See appendix A for an explanation of these.

2.3.3 Specialist specification

The specialist scientist works in a similar fashion to the base agent in that they can add one state to a previous theory. They will be the actual builders of theories going back to their motivation of actively trying to solve problems. However, an added restriction is that they operate in a smaller area and therefore are able to use less previous theories to build onto.

A specialist will theorise a system M' which they hope to be equal to an observed system M'. In their theory creation and analysis they are equal to the base agent. They differ only in what previous theories they can use, which will be discussed below.

2.4 Environment specification

With the formal models of both the system and the agents complete, these parts can be put together to define the environment in which they live. The environment consists of the observed cognitive system, several groups of agents, each with the same number of specialists and generalists, and a theory pool. When all groups have the same agent division, the entire agent population will have this division also. These groups act as the scope limitation for the specialists as they can only select theories to build on from the same group that they are a part of. The generalists can select from all groups (i.e. the entire population). Figure 2.3 shows the environment in full, showing the observed system on top, 12 arbitrary agents observing it, and the theory pool for the theories created by the agents. This way, other agents can find previously created theories, and use them for their new creations. The group structure is maintained within this pool to limit the scientists seeing theories outside of their scope.

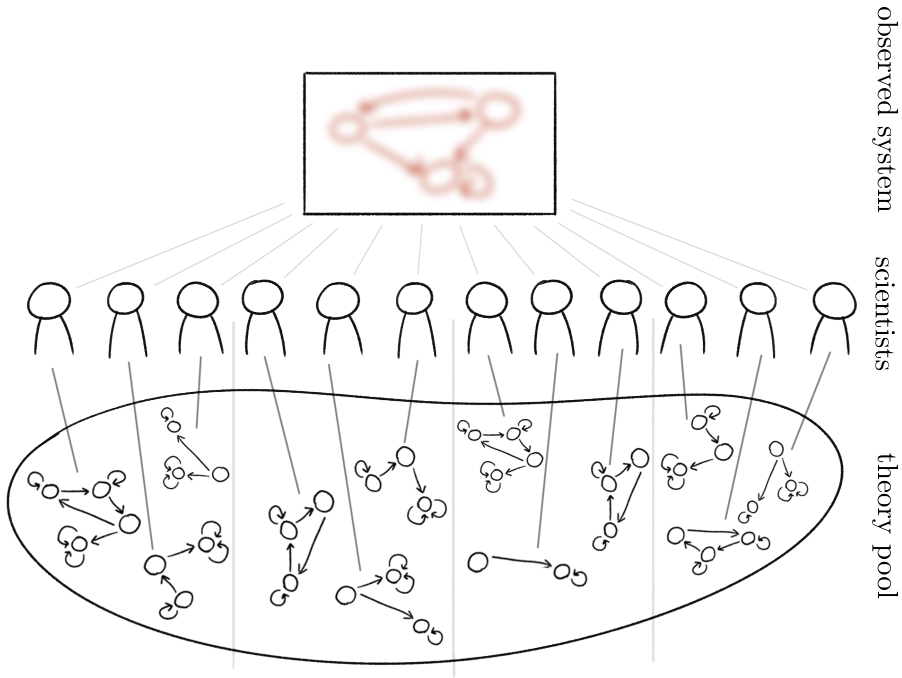


Figure 2.3: The situation for multiple agents. The group division is indicated by the vertical grey lines.

Once either scientist type creates a theory that satisfies them, they will proclaim this to the other scientists in their group. If their satisfaction threshold is the same or lower than the one of the agent that made the satisfactory theory, they will become satisfied also and stop searching. Non-satisfied agents will keep searching.

3 Implementation

Now that the agents have been formally defined as well as the cognitive system they observe it is time to get a bit more practical. The following section discusses how the formal model was transformed into runable code, which can be found online at `https://github.com/mvdmeiden/RU2021-thesis`, an explanation of its structure can be found in appendix B. Firstly, the FST implementation will be discussed and then the agents. Overall, the implementation has been kept as general as possible as long as generalisation did not take an unprecedented amount of time to achieve. This is to accommodate both the project itself as well as possible future research.

3.1 Cognitive system implementation

The FST representing the cognitive system consists of four parameters also seen in the formal mathematical definition: an *input alphabet* Σ, an *output alphabet* Γ, a *starting state* q_0, and a list of *states* Q. On top of that the implementation consists of a *maximum string length*, as well as a list of *empty transitions*. Missing from these parameters are the finishing states F and the transitions Δ. The first have been omitted from the implementation as all states will be finishing states to take away an extra step in implementation. This results in a behaviour where every input results in some output, unless a character is not recognized by the input alphabet. The transitions are embedded in the states: each state has a name and dictionary consisting of the outgoing transitions. In this dictionary, the keys are the characters in the input alphabet and the values are tuples consisting of the next state and a character from the output alphabet. The list of states and their transition dictionaries form the main representation of FSTs. Figure 3.1 shows an example of this list and is a representation of the FST shown in figure 2.2.

There are two new parameters in the implementation that are not present in the mathematical definition. The *maximumstringlength* is to limit the length of the strings inputted and outputted in the observed system. Theoretically, this length is infinite, but for the implementation some value needs to be chosen. The list of empty transitions is needed for the FST generation which will be discussed below. It stores state-input pairs of transitions that are not (yet) part of the FST and it is from this list that new transitions are selected.

In order to use this definition for modelling, the transducers need to be generated

```
q0 {`0': (`q0', `1'), `1': (`q1', `1')}
q1 {`0': (`q1', `0'), `1': (`q1', `0')}
```

Figure 3.1: An example list of states representing an FST. For the state q_0, an input of 0 would result in an output of 1 and the transducer would move to q_0 for the next input.

in such a way that they are both random and have a 'nice' structure. When generating the states and their transitions completely randomly (albeit bound by the parameters above) the resulting structure is not necessarily sound: there might be states that are not reachable from the initial state. This can either be because they have no ingoing transitions or because they have no transitions at all and are completely disconnected (see figure 3.2a for a visual representation). To counter this, the transducers are generated incrementally together with some constraints to always ensure a decent structure. The first of these constraints is that for each state that is added, this state must have at least one ingoing transition. This creates a forward motion in the transducers that makes all states reachable from the initial state — the state that was first added. If the initial state were changed after generation, this would not hold any more. The second constraint is represented by the list of empty transitions mentioned earlier. One of the pairs in this list is selected when a new transition is added. By using this list it is ensured previous transitions can't be overwritten. However, the combination of these two constraints does have some consequences. When there are no empty transitions left, a next state can't be added since no current state is allowed to make an outgoing transition going into our current state, but this transition is necessary according to the first constraint (see figure 3.2b for a visual representation). When this happens the FST generation will terminate and the full FST is returned.

Below, the generation is explained step-by-step:

1. generate the first state which is also the initial state and add the state-input pairs to the list of empty transitions

2. for the next state q, randomly decide how many ingoing and outgoing transitions will be generated.

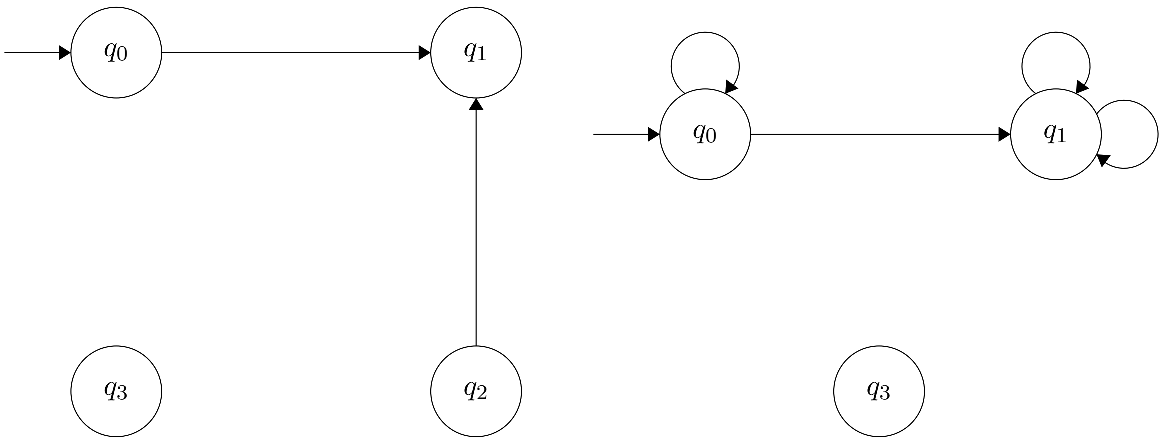
3. for each ingoing transition chose a random pair (q', σ) from the list of empty transitions and create a new transition $\delta = (q', \sigma, \gamma, q)$

4. for each outgoing transition chose a random state q' from all current states and create a new transition $\delta = (q, \sigma, \gamma, q')$

5. add the non-used outgoing transitions to the list of empty transitions

6. repeat step 2-5 until the desired amount of states are reached or the list of empty transitions becomes empty (constraint 2)

This way of generation is used both in M (the system that the agents observe) as well as by the agents to generate theories themselves. Once an FST has been created inputs



(a) An example of an FST of which the structure is undesirable. Here, both state q_2 and q_3 can't be reached from the initial state.

(b) An example where q_3 can't be added to the current FST. This FST has an input alphabet of length 2 which makes it so that there can a maximum of two outgoing transitions per state. q_3 needs an ingoing transition but since no other state can make an outgoing transition and previous transitions can't be overwritten, this transition can't be made.

Figure 3.2: Two problems that can occur during FST generation

can be parsed through it to obtain the outputs, which happens by walking the correct path in the FSTs graph structure. When a certain character occurs in the string but no suitable transition can be walked to the next state, an empty string is returned.

3.2 Agent implementation

Just like in the formal specification, an agent has a *dataset* D of observations and an *acceptance threshold* θ. In addition to this, there is a *complexity limit* that defines how many states the created FSTs can have. An agent can't create nor process theories that are too large, but what happens exactly will be discussed later. First, the basic behaviour of an agent will be discussed.

When an agent chooses a theory from the pool and observes new data, it can act onto this. This acting consists of four steps:

1. create a new theory
2. generate new situation behaviour pairs using this theory
3. compare these pairs with the observed pairs and calculate the accuracy
4. output the new theory, its accuracy, and your scientist type

Only the first step is different for a specialist and a generalist while steps 2 - 4 are the same for both types. A specialist takes a theory and adds one state to its internal structure. It does so in a similar way to generating full FSTs incrementally, but only does one iteration of this incremental process. There are three situations where the procedure differs a bit: when the chosen theory is non-existent (e.g. no theories have been created yet), when the chosen theory is full, and when the chosen theory is too complex. In the first two cases a specialist agent creates a new minimal theory, consisting of two states. In the last case the specialist agent is unable to handle such a complex theory and returns nothing.

A generalist selects two theories from the pool and combines them. They do so by going over the list of states of both theories and for each state select one or the other. When one theory has more states the agent chooses between a state from the larger theory or no state at all. Figure 3.3 illustrates this process. This process isn't perfect as it may create disconnected FSTs in which some states aren't reachable. However, this isn't as big of a problem as it may seem. Sure, these disconnected states can't be reached, but all they do is take up memory space. For the agents it would be as if this state simply isn't there. For the generalist agent there are only two special cases. The case where both theories are full isn't a problem for the generalist as it doesn't need to add any new state. When a generalist encounters a theory that is non-existent it generates a new theory as complex as they can make it. This allows the generalists to create theories when no specialists are around to do so. Instead of this, I could have opted to already add some theories to the theory pool prior to the first iteration, but in the end the choice is arbitrary. Whether I had let the agents create some theories in the first iteration or generate them myself in before that would not have changed the contents of the pool at the first iteration. For the last case, when one or both of the two theories are too complex, the generalists have the same behaviour as the specialists and return nothing.

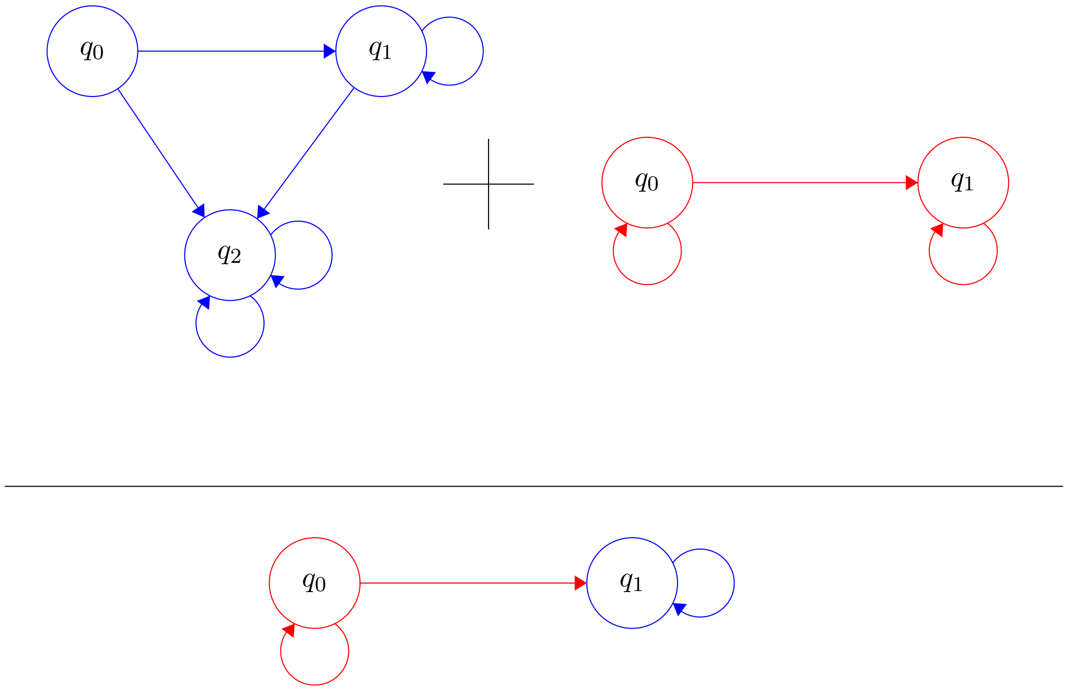


Figure 3.3: The combining of two FSTs. The input and output values have been left out for simplicity and the colours are to indicate what parts were kept.

4 Simulations

In the following chapter, the simulation setup and the simulations themselves will be discussed. Firstly, some preliminary testing will be done to analyse the effects of certain parameters in order to give them a reasonable value for the actual simulations. With these tests complete, the simulations can be set up and run. After the main simulations have been analysed as well, some further testing is done to explore what was learned.
The main testing will be done by varying the distribution of the two agent types over the population. This will hopefully give insight in how each agent type influences the scientific discovery done by the entire population.

4.1 Preliminary testing

There are many parameters that can be altered. Each of these will be discussed and based on that, a value will be chosen for the simulations. Some parameters need no complex reasoning and can be chosen almost arbitrarily, whereas others need a bit more explaining before a reasonable value can be chosen. The first of these two will be listed below:

- number of agents in the population: 100
- alphabet characters: 0 and 1
- acceptance threshold: 0.5
- complexity limit: 50 states

4.1.1 Computational limitations

The most computationally expensive action is the transducing from input to output. This makes that both the number of observations and the length of each observed string are influential in this. When looking at how the dataset of the agents would work, a cumulative dataset was considered. This meant that each agent would add their observations to a shared dataset that would keep growing and a theory would have to explain all data observed up to that point. However, this ever growing dataset would soon get too large as there are a hundred agents in the population each doing observations for many iterations. Instead of this, an agent would do one hundred observations each time they create a new theory. This would keep the memory from getting full and the computations from taking forever. The string length was also kept short, but this was mainly for the reason discussed in the next section.

4.1.2 String length and alphabet size

The maximum string length for the simulations is set to 10. This is not very long, but it is necessary for the agents to reach satisfaction. The agent performance drops as the string length increases as guessing a string becomes harder and harder. For each character, with random guesses, there is a $\frac{1}{|\Sigma|}$ chance to get it right. For a string of length n and an alphabet length of 2, this would come down to a probability of $\frac{1}{2}^n$ to guess the entire string correctly. With shorter strings the agents have a better chance of creating a theory which has an accuracy of at least 0.5. However, this may not be realistic as observed systems often consist of more complex situations and behaviours that can't be represented by a 10 character binary string. It is however also important that some agents create a satisfactory system. If no agent finds anything ever there is no effect to be measured. In order to still be able to observe the differences, this shorter string length was chosen.
Note that I also could have lowered the acceptance threshold to fit with the accuracy of the agents when longer strings are allowed. This would have had the same effect, but would have drastically slowed down my simulations. For the sake of speed, choosing the above option was preferred.

4.2 Setup

Starting a simulation begins with creating the observed system M. This system is made up of a maximum of 50 states, but may consist of just 5 states. The input and output alphabet of all FSTs are the same and consist of the numbers 0 and 1, making them transducers of binary strings. The maximum length of each input and output string is set to 10. Each agent, specialist or generalist, has a complexity limit of two times the size of M and an acceptance threshold of 0.5. Because in the building process theories often become full, not allowing any new states, it would be difficult for an agent to create a theory with the same number of states as M when its size is the agent's complexity limit. This is why this value is set to two times the world size to allow the agents to make more complex theories. The acceptance threshold of 0.5 states that the agents' theories must explain at least 50 percent of the observed data.

A simulation lasts at most a hundred iterations, but can be done earlier when all agents are satisfied. A single iteration consists of one pass over all the agents creating a new theory and adding it too the pool. Each agent observes a new set of a hundred situation-behaviour pairs for each iteration and therefore for each theory that they build.

4.2.1 Data output

Each agent outputs their theory, its accuracy, and their type. This information, together with whether the agent is satisfied or not, is stored inside a table of size 100×100 for each agent and each iteration. When an agent is satisfied and stops research, it will continue outputting their satisfactory theory until the simulation ends. The other agents in their group, which they notified of their success, will also start outputting the satisfactory theory.

4.3 Results

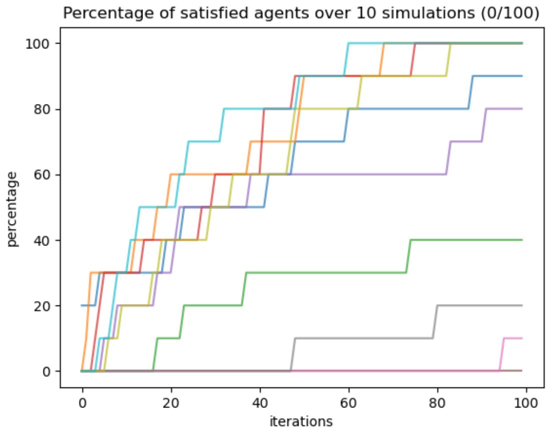
Now that there is a good grasp of how a simulation runs and what affects speed and performance, the actual simulations can be run. Ten distributions were tested and for each distribution ten different simulations were done. This totals out to a hundred different simulations each with their own unique observed system. These runs produced a lot of data and plots and to keep this paper comprehensible I have selected only the most useful ones. Even so, all plots and data can be found online (appendix B).

4.3.1 Satisfaction speed

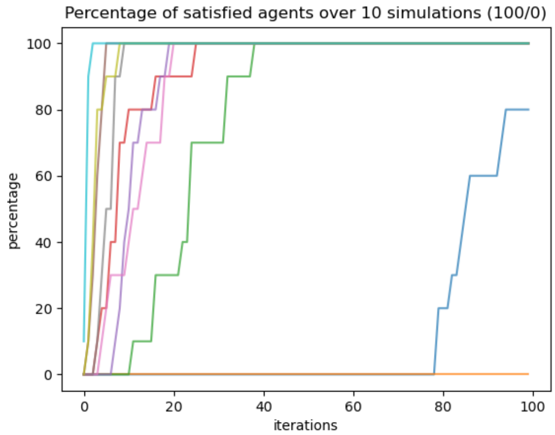
First of all, I wanted to see how fast agents would become satisfied by plotting the number of satisfied agents against the number of iterations. The plots in figure 4.1 show this for several distributions of agents. Each coloured line represents another simulation within the same distribution. Interesting to immediately note, is the large variance in the performance both between simulations with the same distribution and between the different distributions. For an all specialist world however (fig. 4.1a), it takes a lot longer to reach 100% satisfaction as they cannot rely on theories made in other groups. Each group has to find their own theory independently, whereas a distribution with more generalists may find theories from other groups and with that, instigate a ripple effect of discovery. Figure 4.1b seems to indicate the ripple effect is at play as the lines are a lot steeper. This may indicate that as soon as one satisfactory theory has been created, all agents recognize it and become satisfied also. However, it may also be the case the combining tactic of the generalists just creates satisfactory theories faster. What exactly is the cause will be discussed in a later section.

Whatever the case may be, one can at least assume adding more generalist to a population would increase the steepness of these plots. For a distribution of 10 generalists and 90 specialists (from here on out denoted as 10/90) there is not yet any difference in the graph compared to the one with a 0/100 distribution. Once there are 20 generalists in the population a difference becomes apparent (fig. 4.1c). Here, a lot more steep lines are visible but there are still some less heavily slanted lines as well which have completely disappeared when only generalists are in the population. This makes sense as there are still many specialists whose behaviour we see in the less slanted lines. The lines keep becoming steeper and steeper as the number of generalists in the population increases.

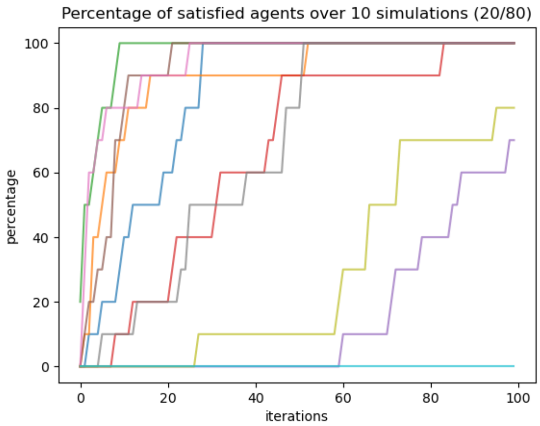
However, even though the satisfaction goes very fast a population with more generalists is a bit of an all-or-nothing situation. With more generalists there were more runs where no agent found a satisfactory theory at all.



(a) 0 generalists; 100 specialists



(b) 100 generalists; 0 specialists



(c) 20 generalists; 80 specialists

Figure 4.1: Several plots showing the number of satisfied agents per iteration. Each plot consists of 10 different runs which are indicated by the different coloured lines.

4.3.2 accuracy variety

For this section I will be taking a look at the end results of each population distribution. After a hundred iterations, what did the agents come up with and how good is it? For the overall final accuracy, only the satisfactory theories were taken into account. Figure 4.2 shows the average accuracy as well as the minimum and maximum found accuracy for each distribution of agents. It shows that there is not much difference in

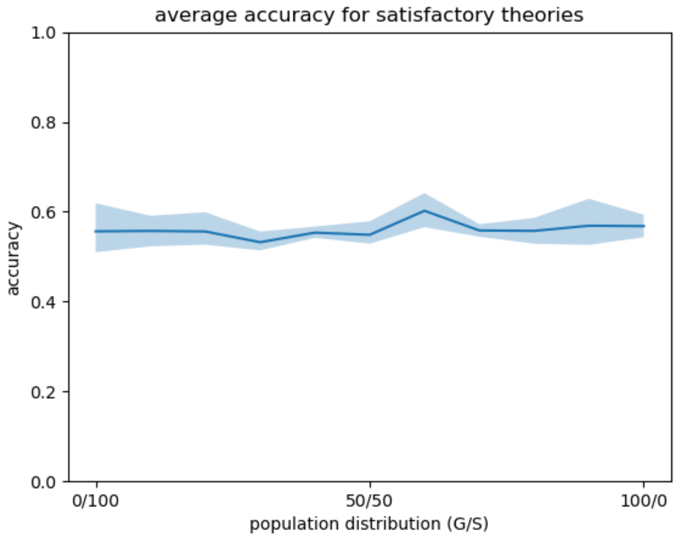


Figure 4.2: The average accuracy and its variance for each population of agents.

quality of theories at all. There is a minor peak at the 60/40 distribution, but there no further indication that this peak is anything significant. This indifference in the results is probably due to the fact that every agent stops searching when they create a theory with an accuracy above 0.5. For this to show significant differences an agent would have to create a theory and jump from below 0.5 to a value far above 0.5. The behaviour of the specialists and generalists are equally random that there is no effect in how far above 0.5 they reach on average. Their different theory building techniques have no influence on this.

So, the variety at the end of each run for all distributions is rather much the same. But how about the variety throughout one run? The agents act randomly and therefore their resulting theories have wildly varying accuracy. There is no structured movement towards a good result and finding a satisfactory theory is just matter of chance. Figure 4.3 shows how the accuracy progresses over time for one run with the 50/50 distribution. All distributions had rather similar plots for this measure, so any distribution could have been chosen to show the effect (or lack thereof in this case). The plot shows how the the accuracy for all theories, where each group has their own colour. When an agent in a group creates a satisfactory theory all other agents get informed about this and the entire group jumps up to the higher accuracy. When another group discovers a this theory they jump to that same accuracy line, but it could also be that they made their own theory with the same accuracy. However, this still gives an indication as to how many distinct theories were created in a run and what groups created what theory. This is also where the most differences were visible between the different populations as there was quite a difference in how many unique theories were created by a certain distribution.

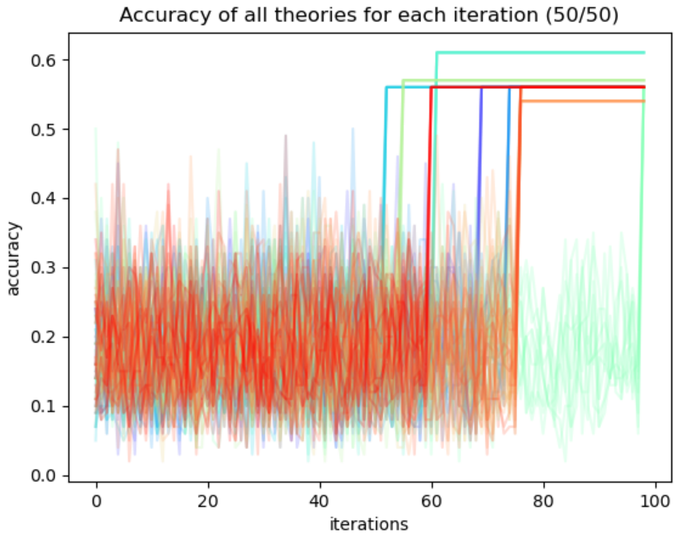
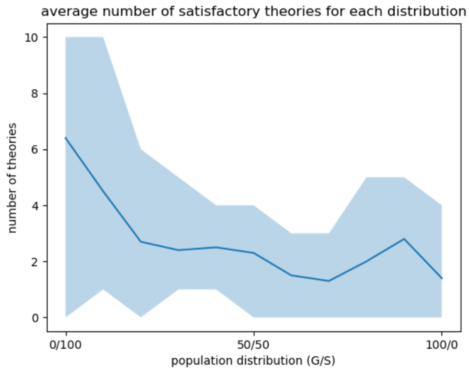


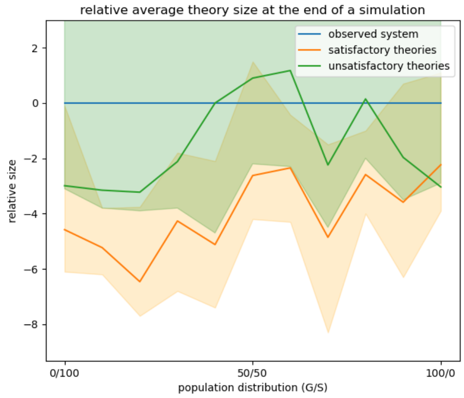
Figure 4.3: The accuracy over a run with 50 generalists and 50 specialists. Each group of 10 agents has its own colour.

4.3.3 Theory variety

The accuracy of the final theories may not have differed a lot, but the number of theories the agents produced did as became apparat when looking at the progression of accuracy over single distributions. Figure 4.4a shows the average number of satisfactory theories found. It shows that, generally, the number of theories found decreases when more generalists enter the population. Just like with the plots in figure 4.1 this becomes apparent at the 20/80 distribution. After this point, the number of theories doesn't decrease as sharply anymore. There is an increase in the number of theories produced for the 80/20 and the 90/10 distributions, but I am unsure what causes this. This measure also gives insight in whether or not generalists actually find theories created in other groups or whether combining theories just creates better theories. If the later would have been the superior theory building strategy, the generalists would also be creating many unique theories, as they would not need to look at other groups to do so. But, since the high-generalist populations only create few satisfactory theories, it can be concluded the ripple effect is utilized and good theories are found in other groups.
Secondly, a look can be taken at the size of theories. The size of a theory is defined by the number of states in that theory and gives a bit of insight in how similar two theories are. For this statistic the distance in the number of states, or the relative size, of the theory to the observed system is what will be looked at. This measure was chosen as an easy option to look at internal similarity between theories. Initially, the idea was to compare the state structure of two theories, but this proved to be unviable as possible isomorphic graphs would be missed. Figure 4.4b shows the size of the observed system



(a) The average number of satisfactory theories at the end of each simulation.



(b) The average size of theories relative to the observed system at 0

Figure 4.4: Two plots showing different aspects of the theories made by agents.

at 0 and two other lines: one for the satisfactory theories and one for the unsatisfactory theories (with an accuracy ≤ 0.5). All of these lines are averages over the ten runs per distribution. Even though the observed system size is set to 0 to act as a base line, the actual average size was 8.6, which may be smaller than initially expected. With a size range from 5 to 50, one would expect the average size to be $(5 + 50)/2 = 27.5$ as the sizes are uniformly distributed. However, due to the FST generation process being halted quite often, this average decreases quite drastically.
Going into the agents' theories it is interesting to see that almost all theories are on average smaller than the observed system. This may be due to the theories only explaining about half of the data rather than all of the data but it may also be the case that the observed system can be described in a simpler manner. For the satisfactory theories, there is an increase in the number of states used as more generalists are coming into the population implying that specialists in general need less states to reach satisfaction. What could play a role here, is that specialists start with a theory of two states and build onto it whereas generalists start by creating a theory as big as they possible can. However, the average of the generalists is still well below the size of the observed system even though their capabilities are twice that.
The plot for the unsatisfactory theories makes a peculiar shape. They are generally a bit larger than theories with a higher accuracy but there is no positive or negative trend in the line. Ignoring the dip at the 70/30 distribution, a bell shape can been seen in this line which would mean there is something about the distributions with 50-80 generalists and 50-20 specialists that can create large theories. It is hard to say what this something is, but it could be the interplay between the generalists breaking down the incremental structure build by the specialists and with that opening up enough space for the specialists to keep on building without running into too many full FSTs.

So far, I've only compared different populations with each other. When zooming it a bit, plots like those in figure 4.5 emerge. Here, the size of each theory made for the ten runs with one distribution is plotted. This shows the progression of theory size where the darker areas show within what range most theories were produced. With the all specialist population (fig 4.5a) we can see the theories increase in size gradually as the simulation progresses. This is fully expected as specialists build their theories from the ground up. With generalists it is exactly the opposite (fig 4.5b): their theories start large and become smaller, converging to somewhere between 4 and 6 states, kind of like a sideways funnel. Generalists can't make theories larger than the ones they select from the theory pool. Their resulting theories lay somewhere between the sizes of the two selected ones, but will never exceed them. Because of this, the resulting plot shows the expected behaviour.

The specialists tendency to make increasingly larger theories is a strong one. Even when there are only 10 specialists present in the population the graph starts displaying a small upward curve towards the end (fig 4.5c). This upward motion stays apparent as more and more specialists enter the population, but doesn't necessarily become more steep. It just becomes more prominent and less and less large theories are made in early iterations.

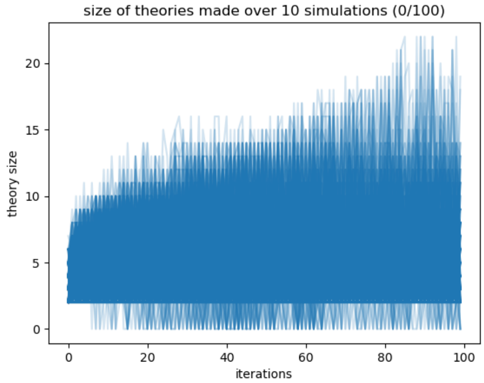
4.4 Further testing

4.4.1 Performance against chance

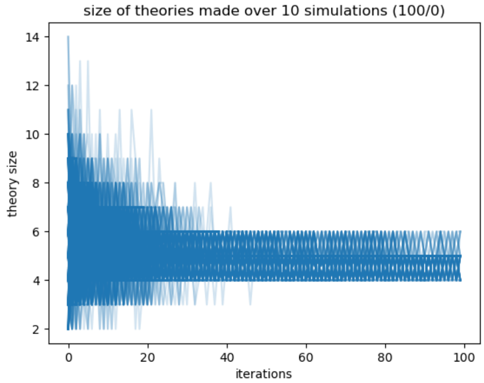
While discovering the effect string length had on performance I figured it would be interesting to test the agents against a chance baseline. For this, I ran a couple simulations where the acceptance threshold was set to 1 — so no agent would ever be satisfied — and varied the maximum string length. The plot in figure 4.6 shows three lines: the base accuracy which is based on random chance, the average agent accuracy for a 50/50 distribution, and the average highest found accuracy. The average was taken over 5 simulations and the base accuracy was calculated according to the function below:

$$P(\text{correct guess}) = \frac{1}{n+1} \cdot \sum_{k=0}^{n} 0.5^k$$

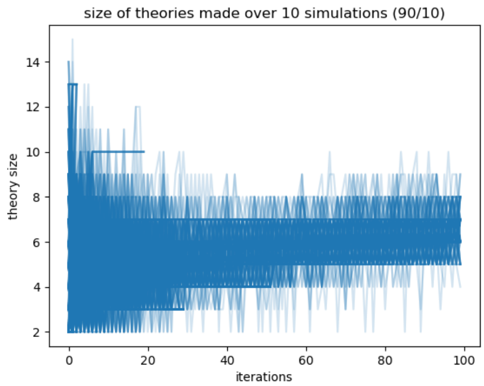
This function takes the average of the possibilities of guessing a string correctly for every possible string length where n is the maximum string length.
The graph shows that the agents actually perform worse than chance on average, but it also shows there is a large difference between the highest accuracy and the average. This leads to the interest of the entire distribution rather than just the average. Figure 4.7 shows how often each accuracy value occurred in 5 runs with a maximum string length of 10. Most accuracies were between 0.1 and 0.2 and the distribution seems to resemble a Poisson distribution. There are very few accuracies above 0.5 and even fewer above 0.6 (they're not visible in the graph anymore). However, these few values are enough



(a) 0 generalists; 100 specialists



(b) 100 generalists; 0 specialists



(c) 90 generalists; 10 specialists

Figure 4.5: theory size

for a proper simulation as finding one satisfactory theory can bring about a ripple effect allowing it to be found by agents in other groups.

4.5 In summary

From the simulations it became clear that the accuracy of the FSTs created by the agents was largely dependent on the string length of the input. Having a more realistic, long, string length meant the agents would have a very slim probability of creating a theory, which accuracy would be large enough to satisfy them. As the progress in this model is purely dependent on how many agents are satisfied, making it nearly impossible for the agents to reach satisfaction would just result in no effect to be seen. To allow the agents to at least sometimes run into a satisfactory theory, the string length was lowered. Because of this though, any effect that was seen after this, only occurs in a situation where agents can actually come up with a reasonable theory.

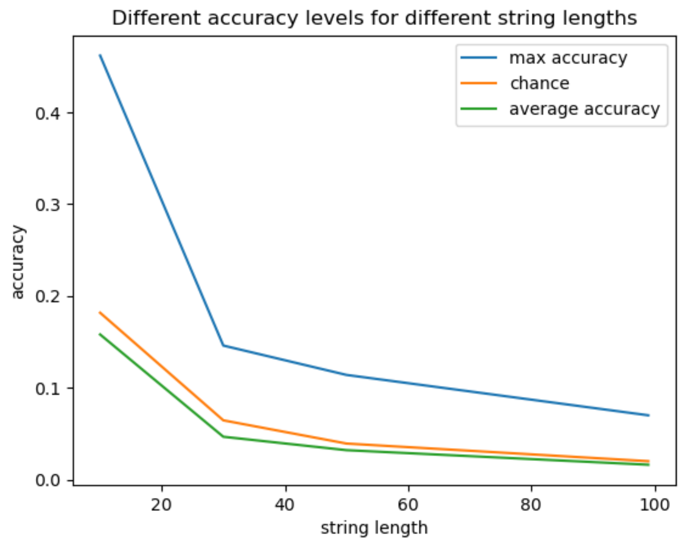


Figure 4.6: A graph showing different average accuracies for different string lengths. The averages were taken over five runs.

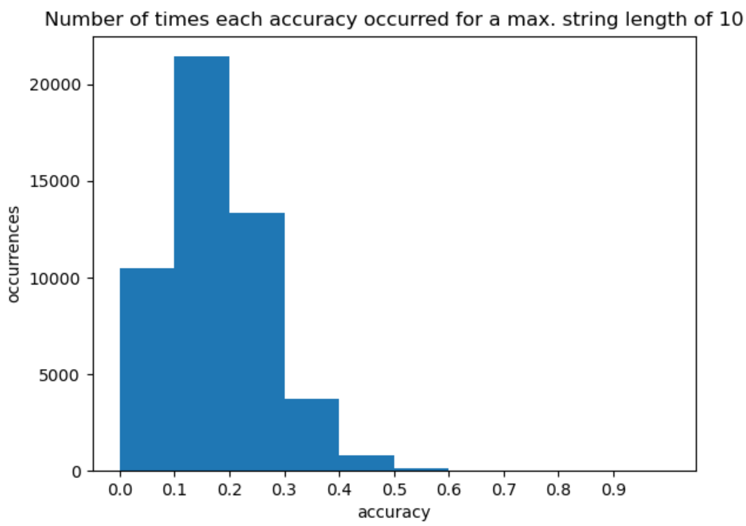


Figure 4.7: the number of times a certain accuracy occurred over 50.000 samples. The accuracies were divided into ten groups based on the first decimal number.

With this in mind, the actual simulations were run and differences between the two agents were observed. To measure the progression of scientific discovery the speed with which different populations of scientists became satisfied was measured. These measurements varied greatly from run to run and sometimes there were large stretches of time between the discovery of two satisfactory theories. Overall, populations with more generalist agents were able to reach satisfaction faster, but also had more runs where no satisfactory theory was found at all. After looking at the number of unique theories created by each population, it could be concluded that the generalist agents were using their ability to find satisfactory theories within other groups to become satisfied so quickly, rather than making multiple satisfactory theories themselves.

Next, the created theories were compared to the observed system, to see how similar they were. In the first instance, the idea was to compare the internal structure of the finite state machines. However, this proved to be an unuseful metric as possible isomorphic graphs would be missed. Instead, I decided to look at the number of states the FSTs consisted of relative to the number of states in the observed system. Overall, theories of all sizes were created, both smaller and larger than the observed system. On average however, most theories were smaller than the observed system and satisfactory theories were smaller than non-satisfactory theories. When looking at the differences in size between different populations, it was less clear what could be observed from this. The difference between the observed system and the satisfactory agent theories, became gradually smaller as more generalist agents joined the population. Specialist-heavy populations on the other hand made smaller theories, with a bigger difference to the observed system. Since all of these theories are satisfactory to the agents, one can wonder which is better: to be more similar in size to the observed system but larger, or to be further away from the observed system but smaller. Personally, I would argue the smaller the theory, the better, especially since they only explain a bit more than 50% of the data. Therefore they perhaps shouldn't be that close to the observed system in size at all. A smaller theory leaves more room for improvement and is overall easier to interpret.

Besides looking at theory size between the distributions, it was also measured between different runs for one distribution. This showed how the theory size progressed over a run for each distribution. As expected, the theory size of the specialist agents increased slowly as they kept adding states, creating larger and larger theories. The generalist agents start of with theories of all sizes and then quickly converges to a stable size. The specialists here have a large effect on the theory size, as even with just a few specialists in a population, theory size increases, rather than stay stable.

When looking at the agents overall performance they were measured against a chance baseline. This baseline was based on the chance to correctly guess a string. As it turned out, the agents performed worse than chance on average, but the large variety within their guesses still allowed them to find a satisfactory theory often enough. As to why they perform worse than chance, if only a little, is hard for me to say. When creating the internal structure of a theory, rather than just guessing the resulting behaviour of a certain situation, a lot more variables are involved that determine your success. And even if constructing theories may be worse than chance at times, random guesses will never lead to an understanding of the inner workings of a system. Therefore they will not accomplish anything, even if the results may seem more preferable at first.

5 discussion

In the above chapter, the differences between the agents within the model have been analysed. Out of all the differences, some were artefacts of implementational choices rather than conceptual ones. For example, the progression of theory size for both agent types is dependent on the size of the starting theories. These starting points were chosen arbitrarily and not because they highlight any differences between specialists and generalists. Of the differences that were the result of conceptual choices it was nice to see the generalist agents actually use the ripple effect.

Only the distribution of the agent types was manipulated and most other variables were left unchanged. Of these variables, most were left alone because either their effect would be obvious or computation power didn't allow it. The complexity limit on the agents' theories however is an interesting one. Although it was set high enough for the agents to create large complex theories, this almost never happened because of the way these were generated. The created theories never reached a large size and therefore I wasn't able to mess around with the parameter either. Although I suspect nothing much would have changed if the systems were able to larger except for maybe an increase in the time it would take to reach satisfaction. However, it is a shame this could not be tested because of the way the FSTs were generated and using different ways to generate FSTs might be an interesting continuation of this research.

Besides the unchanged parameters, there were some parts in the agents that lacked proper motivation and were more of a consequence of implementation. The behaviour of each agent when no theories were created yet is one of these things. To allow the generalists to have some theories to work with in a population with no theory building specialists, I allowed them to create theories themselves in the first iteration. This obviously influenced the shape of the size progression plots in the previous chapter.

Besides this, there is one major part I simply forgot to implement. In the current setup, the generalist agents observe an equal amount of data as the specialists. This gives them the same amount of knowledge, whereas earlier I argued that generalists only have surface-level knowledge and one way to implement this, would have been to limit the amount of data they observe. I suspect my focus on the different theory building strategies obstructed away from a perhaps trivial implementation. But what could have happened if I had implemented it? One might argue that the generalists would have performed worse as they'd have less data to work with. On the other hand, they'd have less data to explain and may have gotten satisfactory theories even faster. Because of the random data generation and the random behaviour of the agents, I don't think the amount of data the agents observe is a proper aspect of the agent implementation anyway — within my model that is.

This paper started out with formal models of a base agent, a generalist agent and a specialist agent. These last two were based on previous research and their most defining aspects were chosen to be incorporated in their model. Although some aspects were missed, the implementation was then used to run simulations with different population distributions. Having run all the tests and analysing what effects there were. Within the simulations and within the model a difference between the two scientist types is visible and even though some of these differences are only due to implementational choices and not underlying differences, it is still nice to see some differences arise. Even so, no single scientist distribution is marginally better than the other and each have their own advantages. So, are there differences between different distributions of generalist and specialist scientists? Yes. Do these differences have any monumental effect or consequences? Not so much. Science is just very difficult and whatever distribution of scientists you have isn't going to make science super easy all of a sudden. Coming back to Bateman's statement about how being a certain scientist 'type' is more of a choice in tactics rather than something one either is or isn't, it is entirely up to the individual to become what they see fit.

Bibliography

Bateman, T. S., & Hess, A. M. (2015). Different personal propensities among scientists relate to deeper vs. broader knowledge contributions. *Proceedings of the National Academy of Sciences*, *112*(12), 3653–3658. https://doi.org/10.1073/pnas.1421286112

Berstel, J. (1979). *Transductions and context-free languages*. Vieweg+Teubner Verlag. https://doi.org/10.1007/978-3-663-09367-1

Bode, H., Mosteller, F., Tukey, J., & Winsor, C. (1949). The education of a scientific generalist. *Science*, *109*(2840), 553–558. http://www.jstor.org/stable/1676674

Devezer, B., Nardin, L. G., Baumgaertner, B., & Buzbas, E. O. (2019). Scientific discovery in a model-centric framework: Reproducibility, innovation, and epistemic diversity. *PLOS ONE*, *14*, e0216125. https://doi.org/10.1371/journal.pone.0216125

Payette, N. (2012). *Agent-based models of science*. https://doi.org/10.1007/978-3-642-23068-4_4

Rice, S. A., Wilson, F. G., & Cook, T. I. (1950). Generalists versus specialists in social science: Discussion. *The American Political Science Review*, *44*(2), 380–393. http://www.jstor.org/stable/1950277

Rich, P., Haan, R. D., & Wareham, T. (2021). *How hard is cognitive science?* https://doi.org/https://doi.org/10.31234/osf.io/k79nv

Teodoridis, F., Bikard, M., & Vakili, K. (2019). Creativity at the knowledge frontier: The impact of specialization in fast- and slow-paced domains. *Administrative Science Quarterly*, *64*, 894–927. https://doi.org/10.1177/0001839218793384

Wray, K. B. (2005). Rethinking scientific specialization. *Social Studies of Science*, *35*(1), 151–164. http://www.jstor.org/stable/25046633

A Multisets

Sets are defined in such a way that each element may only occur once. However, when defining the generalist agent elements need to be able to occur more than once. For this, multisets are used. In a multiset each item has a multiplicity: the amount of times an item occurs. Multiset addition is denoted with $\uplus$ wherein the multiplicities of each element are added together. See the example below for clarification:

$$\{2, 2, 3, 3, 3, 5, 8, 8\} \uplus \{2, 4, 5, 5, 8, 8\} = \{2, 2, 2, 3, 3, 3, 4, 5, 5, 5, 8, 8, 8, 8\}$$

an alternative notation of the line above where the multiplicities are clearly specified:

$$\{2 : 2, 3 : 3, 5 : 1, 8 : 2\} \uplus \{2 : 1, 4 : 1, 5 : 2, 8 : 2\} = \{2 : 3, 3 : 3, 4 : 1, 5 : 3, 8 : 4\}$$

B Online resources

The code for the paper can be found online at `https://github.com/mvdmeiden/RU2021-thesis`. In the following section, its content and structure will be explained. The GitHub consists of three main folders: code, data, and imgs.

code

The first one contains all the code used during the simulations and one additional file. This is the `MealyMachine.py` and contains the base code used to implement the FSTs.

data

The second folder, the data, contains all generated data from the simulations.

imgs

This final folder contains all images used in the paper as well as plots of the data that were not shown, but can be viewed if you so desire.