Urban Planning-in-Pieces: A Computational Approach to Understanding Conceptual Change and Causal Reasoning about Urban Planning

Arthur Hjorth, *arthur.hjorth@u.northwestern.edu*
Center for Connected Learning and Computational Modeling, Northwestern University, USA

Uri Wilensky, *uri@northwestern.edu*
Center for Connected Learning and Computational Modeling, Northwestern University, USA

Abstract

In this paper, we present a computer-simulation-based classroom unit designed to improve students' *causal* reasoning about the effects of urban planning. Based on data collected in three separate implementations of this unit, we use Association Rule Mining (ARM) to assist in Knowledge Analysis (KA) of students' responses to pre- and post-questions. We first define *causal-nodes* as a construct, and then qualitatively identify distinct causal-nodes in students' responses. We then validate these nodes quantitatively within and between students' responses. Finally, we compare changes in the association rules between these causal-nodes between pre- and post-responses in order to find changes in students' explanations. This paper makes two distinct contributions: first, it shows a productive use of an existing computational method to aid in the analysis of conceptual change. Second, it contributes towards better understanding how to design for and analyse causal reasoning in social science education.

Keywords

knowledge analysis; netlogo; social studies; complex systems; causal reasoning; computational analysis; computational methods

Causal Reasoning in the Social Studies

The College, Career, and Civic Life (C3) Framework for Social Studies State Standards (NCSS, 2013) presents a new framework for K-12 social studies, emphasizing 'student explanations' and 'complex causal reasoning' in Social Studies. However, in contrast to a similar move in science education (NGSS Lead States, 2013), design and implementation of *social studies* curricular activities that focus on causal reasoning are presently understudied. This leaves two gaps that we think are important in the literature: first, how do we design to encourage and strengthen causal reasoning in social studies; and second, how do we study this thinking in a social studies context?

In this paper we hope to contribute to these currently understudied goal by presenting a computationally-assisted Knowledge Analysis (diSessa & Sherin, 2015) of students' causal-explanations in response to pre- and post-questions as part of a unit of our own design, in which students used a computer simulation to learn about urban planning. The simulation and accompanying activities were designed to help students think *causally* about the social impact of urban planning decisions. Analysing students pre- and post-question responses, we identify a set of *causal-nodes* and use Association Rule Mining (ARM) (Agrawal, Imieliński, & Swami, 1993) to look at how, at the classroom-level, these nodes are reorganized between pre- and post-responses. Based on our findings and discussion, we claim that ARM is a particularly well-suited approach to helping us in the analysis of student reasoning with a piecemeal view of knowledge and learning, and that it potentially can help us scale up Knowledge Analysis to larger dataset than it has currently been applied to.

Research Questions

In this paper, we address the following questions,

1. How can we design activities that improve students complex causal reasoning about social issues?
2. What are the *constituent parts* of students' causal-explanations when responding to questions about city planning?
3. How are these parts assembled into larger explanation-structures, and how can ARM help us make sense of change in these explanatory structures?

Why Association Rule Mining for Knowledge Analysis?

Knowledge Analysis takes a piecemeal approach to understanding the mental representations of knowledge, and views reasoning as the *assembly* of these pieces in-the-moment (diSessa, 2002; diSessa & Sherin, 2015; diSessa & Sherin, 1998; Sherin, 2006; Sherin, Krakowski, & Lee, 2012). The purpose of KA is to provide an analytical space for better understanding knowledge and learning: In this view, briefly, learning is the acquisition of new knowledge pieces into a learners' repertoire, and/or changes in the assembled constellations of these pieces. An important feature of a single knowledge piece, is that it can be combined with other knowledge pieces into essentially different, larger structures, and that the meaning of a network of knowledge pieces is an emergent property of these manifold structures.

Likely due to the fine-grained nature of KA, most KA-oriented studies are based on somewhat small sample sizes. Recent work has moved to include more computational methods in the Learning Sciences (Martin & Sherin, 2013; Sherin, 2013). In contrast to Educational Data Mining's focus on quantitatively *assessing* student thinking, this work has primarily focused on using computational methods to qualitatively better understand the process of learning or the constituent parts of conceptual change (Berland, Baker, & Blikstein, 2014; Blikstein, 2011), and on utilizing the power of computation to scale up the size of data. In this paper, we propose to use Association Rule Mining for a similar purpose: as a method for assisting us in a Knowledge Analysis, both as an approach to validating the causal-nodes that we identify in students, and as a way of scaling up KA to help us look at learning at the classroom-level.

Association Rule Mining (Agrawal et al., 1993) is a method for understanding co-occurrences between elements in a set of data. More specifically, an ARM-analysis takes a set of 'transactions' that each include some elements, and then calculates, across all these transactions, how well the presence of one particular element in a transaction predicts the presence of another element in a transaction. We believe that ARM's focus on looking at the co-occurrence of individual items makes it particularly well-suited for helping us in analyses of students thinking with a piecemeal view of knowledge: If we view student whole explanations as transactions, and knowledge-pieces as the individual elements in these transactions, we can use ARM to better understand the changes in students' reasoning both at the individual- and at the classroom-level by calculating if and how these pieces are reassembled over time.

Previous use of simulation in urban planning education

Simulations have been used for SimCity in particular has received attention since the early-1990s as a teaching tool in formal education that involves urban planning or thinking about cities (Adams, 1998; Dorn, 1989; Gaber, 2007; Kolson, 1996; Pahl, 1991). Early uses of SimCity focused on two different aspects. The first sought to make a general case for using commercial computer simulations games in education (Shaffer, Squire, Halverson, & Gee, 2004; Shaffer, 2006; Shaffer, 2006; Squire, 2003) and sought to show how the well-designed and engaging commercial games would make school fun for children. The other focused on simulation literacy (Gee, 2003, 2007; Turkle, 1997), and its specific interest was in how (or if) children engaged with the *ideological* assumptions programmed into the city. More recent work has focused using SimCity in formal *urban planning* education at the college or graduate level: how to align SimCity with formal curriculum and use it as an introductory tools at the college and graduate level (Bereitschaft, 2016; Devisch, 2008); how SimCity can help urban planning students engage with their own creativity (Kim & Shin, 2016); or as a reflective tool to improve on urban planning pedagogy itself (Kim & Shin, 2016).

However, while simulation of *causal* mechanisms are at the very core of the representational power of simulation games (Frasca, 2001, 2003; Squire, 2003), none of the reviewed work has focused specifically on causal reasoning. Thus, while simulations in general and SimCity in particular have seen use in education, this work does not address the specific need created by NCSS (2013) for a conceptual framework for analysing and measuring learners' complex *causal* reasoning. One aim and contribution of this paper is to fill out this gap in the literature and explore how to design simulations to help students improve on their complex causal reasoning about complex social phenomena. We previously reported on qualitative findings from pilot data from this study (Hjorth & Wilensky, 2014a, 2014b; Hjorth & Krist, 2016), but have now collected enough data to take quantitative approaches like this as well.

The Design & Study

The data in this paper come from a unit that we designed on Urban Development and Regional Planning and which we implemented during two quarters at the undergraduate level at a mid-sized, private research university in a metropolitan area in the American Midwest. The course was called, 'Introduction to Social Policy' and is required for Social Policy majors. Students were given course credit equivalent to one course essay for participating. In this paper, we focus on students' pre- and post-responses to a question about how urban planning affects the distribution of commute times for different income groups. Over the span of three implementations, a total of 60 students consented. We sent out the pre-questionnaire 10 days before class in which we used the model, and we sent out the post-questionnaire 10 days after class had finished, and students typically responded within two days. Not all students responded to our questions, and we had some attrition between pre- and post-, and *during* students' responding to questionnaires due to technical problems. In the end, we had 41 students who responded to both the pre- and the post-question.

Our Design and Activity

We designed a unit that ran over the span of two class periods in which students use a NetLogo (Wilensky, 1999) simulation to build cities. Before the simulation activity, the professor (who was not part of the research team) first led a 45-minute discussion about *why people live where they live*. The purpose of this discussion was to cue students' causal reasoning about the role of human decisions, and how these decisions play a causal role in the emergence of economically segregated neighborhoods. Students then worked together in groups of three over a period of about an hour and a half, using the simulation.

The simulation and activity were designed to help students iteratively improve their causal-understanding. We did this by letting students iteratively *articulate* a causal-explanation, *test* it, and potentially *revise* it. This was achieved by designing an iterative, four-phase activity:

1. The Design Phase

The primary purpose of the design phase was to prompt students to reason causally about how to design a city that meets a set of measurable policy goals. Students were first asked to write a set goals for their city. This included specifying which one of three policy outcomes – commute times, local neighbourhood school funding, and access to parks & leisure areas – to focus on, and setting measurable goals for that outcome. Examples include, "We want everybody to have less than 30-minute commute time", or "We want the poorest 20% to have as much funding for schools as the wealthiest 20%". They would then be asked to describe how they would achieve these goals. Students responded along the lines of, "We will put highways everywhere so there are enough roads for everybody", or "We will put parks all over the city to make it attractive for wealthy people to live anywhere so their property taxes are spread out across the city." (Schools are primarily funded through property taxes in most of the US.)

2. The Implementation Phase

During the Implantation Phase, students built their cities inside the simulation. The simulation is designed to let students do this in various ways: Students designate zoning in the city, specifying the density of dwellings, and making certain areas more or less desirable and relatedly more or less expensive to live in. They can also put parks in the city, or build railroads or highways for people to commute on. Finally, they designate certain parts of the city as 'Industrial areas' where jobs will be located.

3. The Growing Phase

The purpose of the Growing Phase was to let students see the *dynamic* effects of their design decisions. During the Growing Phase, the computer model simulates and visualize how (AI-based) computer-agents move into the city that students built. These agents have different income levels and different job locations and make decisions about where to live based on affordability, desirability, and distance to their job. When agents move into an area, their income is reflected in the house prices of nearby areas, so if a wealthy person moves into a neighbourhood, house prices go up, and vice versa. Agents also use the roads near them to go to work, and the more agents use a road, the more congested it gets, making nearby areas less attractive.

4. The Data Analysis Phase

The purpose of the Data Analysis phase was two-fold: first, it was for students to use data as a means of assessing the success of their city; and second, for students to potentially revise their causal-understanding of the model in the cases when their cities did not meet their policy goals. There are two different ways in which data can be visualized in the model: one is spatial, and the other is with bar charts. They provide different perspectives on the same questions, but it is often necessary to look at both in order to really make sense of how the city evolved. Bar charts showed how each income decile was affected by each of the three policy outcomes measures, and the map helped students visualize the geographic distribution of how people were impacted by the policy outcomes.

Data & Analysis

The student responses that we focus on in this paper were all in response to the question,

"Can you explain why a wealthy person's income might make their commute time longer than a poor person's?"

A conventional explanation, consistent with real-world data, could include a variety of factors and sound something along the lines of, "In American cities, higher-income people often live in suburbs, because they can afford to buy houses there, and because they can afford to own cars that allow them to commute between their workplace in the city and the suburbs that often have no public transit options. They choose to do so because they want to live in places that they perceive as safe, and have well-funded public schools. High paying jobs are often located in the downtown areas of cities, and consequently high-income people must make the commute in and out, often on congested roads due to the number of people who also commute in and out at the same time". While it is true that some high-income people live closer to their jobs than some low-income people – in part because they can afford more freedom to choose where to live, and where to work – the result is nonetheless that higher income people typically have longer commutes, but with a large spread in a bimodal distribution. We were curious about how students would reason about it exactly because it contrasts many people's initial assumption that a higher income leads to more desirable outcomes by all measures.

While some KA-approaches have strict, conceptual selection criteria for the pieces they identify (diSessa, 1993), in this study, we take a pragmatic and somewhat promiscuous approach to identifying the individual knowledge-pieces. Rather than looking for a particular grain size or ontogenetic origin, we used 'causality' as a sensitizing concept (Miles & Huberman, 1994) when identifying knowledge-pieces in students' explanations, and looked for parts of their explanation that we could put "because… " in front of. Because our knowledge-pieces relate to students' reasoning about causality, and inspired by Sherin, Krakowski and Lee's (Sherin et al., 2012) *node-mode* approach towards a more permissive inclusion of students' knowledge-pieces, in the following, we will refer to them as 'causal-nodes'. Further, we will refer to the process of *using* a casual-node when constructing an explanation as 'activating' that causal-node, and we call the process of combining causal-nodes into larger explanations as 'co-activating' those causal-nodes.

Causal-Nodes and Student Response Examples

To give the reader a sense of what causal-nodes in our data look like, we provide some examples of student responses and connect them to our causal-nodes. The student responses are always reproduced in full, and as the student wrote them. Across all 82 responses (41 pre- and post-responses) our first round of coding identified 21 different causal-nodes. We iteratively condensed this set twice, eliminating similar codes, eventually arriving at a total of 9 different causal-nodes. Table 1 provides a description of all nine causal-nodes, grouped by what we think of as three interesting types: the first relates to the *geographic location* of people and their jobs. The next group relates to how a person's income affects their *actual commute* – either by influencing their available modes of transportation, or things that affect their commute speed. The final type relates to *the wants or desires of people*, and how having money better allows high income people to fulfill them.

As Table 1 (next page) shows, the most frequent causal-nodes were the ones that deal with the geographic location of people and jobs. Consider the following response,

A wealthy person has greater freedom to choose where to live and often wealthier neighborhoods are in suburbs away from urban areas. (S2-post)

In this response, we see two different causal-nodes: first, that money gives people more choice, and second, that suburban areas in which wealthy people choose to live are geographically far from urban areas. However, the response does not explicate exactly how these two causal-nodes end up resulting in longer commute times for higher income people. In contrast, consider this elaborate student response,

> *Often wealthy people live in a suburb outside of the city because they can afford a house out there and the schools are often better. They also often work in the city at a job that is paying them enough to be able to afford to live in the suburbs. Another big thing is that more wealthy people can afford a car or are able to pay for the train everyday so that they can live far away from their work. (S7-post)*

This response contains the same two causal-nodes as the previous response: people with more money have more choice, and they choose to live in suburbs. However, we see additional causal-nodes in this response: first, the response explicitly states that their jobs can be in the city; second, that this causes them to have to travel a larger geographical distance; and finally, that owning a car is expensive and that wealthy people can afford to own one (or afford to take commuter trains.) These were somewhat typical responses, but they show how he different constellations of casual-nodes can result in different responses, or in similar kinds of responses with variations in what students focus on. We also saw responses that focused more on practical issues relating to the process of commuting,

> *A wealthy person probably commutes by driving his or her own car. A person in a car is subject to traffic from stoplights and other vehicles. Trains, on the other hand, can travel at a quicker rate and also don't have to stop at lights or wait for other vehicles. (S37-post)*

This response does not activate any of the geographic location-related nodes, but focuses purely on the commute-related ones. However, we also saw some responses that seemed to activate different causal-nodes that included perceived differences between *why* wealthy people choose to do what they do, even in the face of a longer commute time,

> *If they choose to live in a suburb or nice area that is farther for work because of the neighborhood or home or school system. They also more likely have ways to commute comfortably and efficiently. (S23-post)*

Students that reasoned about *why* wealthy people make the choices that they do often focused on their perceived better school systems, more green areas or nicer houses in suburbs. But we also saw a few responses that essentialized different characteristics or desires in low- and high-income people. For instance, in this response,

> *They might have more time to spend on leisure, and might not be in a rush. Since they make enough money, they don't have to work as much. (S10-pre)*

Node_ID	Causal-Nodes ('Because…')	Pre-frequency	Post-frequency	Within-person stability	Combinability
Group 1: Geographic/Spatial Location					
0	High income people live in suburbs / low income people live in the cities	0.73	0.83	1.13	8 / 8
1	Jobs are located in cities, not in suburbs	0.46	0.49	1.19	8 / 8
2	High income people may live further from workplaces / Low income people might live closer to workplaces	0.63	0.59	1.12	8 / 8
Group 2: Commute-Related					
3	Buying/owning a car is expensive	0.1	0.1	5.12	7 / 8
4	High income people more likely to own and commute by car	0.1	0.05	5.12	5 / 8
5	Expressways/highways can be congested, or driving can be slow	0.34	0.2	2.56	7 / 8
Group 3: Desires or Possibilities relating to Income					
6	High income people have more choice / low-income people have less choice	0.59	0.59	1.14	8 / 8
7	Wealthy people care less about commute time / are willing to commute longer	0.12	0.07	2.73	6 / 8
8	People want to live in safe areas or places with more space or green areas or better schools or nicer houses	0.39	0.34	1.1	7 / 8

Table 1. Causal-nodes and descriptive statistics

the student seems to construct an explanation around an assumption that higher income people somehow don't feel the same sense of urgency, or that they have a different set of preferences around their choices when it comes to commute times.

As the examples above hopefully illustrate, we saw a large variety in the types of student responses, and in the ways in which we see them activate and co-activate various causal-nodes. Some student responses only included causal-nodes from one or two of the three causal-node groups, while others mixed them across groups.

Validation: Are these Knowledge-Pieces?

KA views knowledge as constructed in the moment through the assembly of knowledge pieces. However, it does of course not view this process as random. Rather, taking a Knowledge Analysis approach to making sense of student reasoning, we would expect to see three different properties relating to the consistency of the nodes that we identify:

The pieces should be found somewhat frequently in a specific population that reasons about a given domain or question (i.e. 'frequency *between people'*). This would indicate that these are more general thinking-bits, and not only parts of one person's idiosyncratic way of thinking.

The pieces should be generatively combinable, meaning they should be combinable into different kinds of explanations (i.e. *frequency between different explanations*). This would indicate that they are truly 'pieces', and not in themselves larger explanations. Finally,

Even though individuals might acquire new nodes and/or reorganize their existing ones, we would expect there to be *some degree of stability* in the knowledge-nodes that students activate when responding to a similar question in a short timeframe (i.e. 'frequency *within* people'). This would indicate that these are parts of somewhat stable knowledge structures, and not completely randomly selected when the student is prompted to answer a question.

Validating Frequency Between People

This criterion is straight forward to validate. We calculated the frequency of each node in pre- and post-responses. The results can be seen in **Error! Reference source not found.** in respectively pre- and post-frequency. We see that even the least frequent of our codes appear in at least 10% of responses in either pre- or post. While the exact cut off for this rule is contestable, we believe that seeing the activation of a causal-node at some point in time across 10% of responses seems like a reasonable number.

Validating Frequency between Different Explanations

There are two different ways in which this can be validated. First, a very simple quantitative statistic showing with how many of the other 8 causal-nodes that we see each node co-occur. In the 'combinability'-column in **Error! Reference source not found.**, we show how many *other* causal-nodes we see each causal-node co-occur with across all responses. Even the least frequent causal-nodes (3 and 4) are used respectively with 7 and 5 of the other 8 casual-nodes, suggesting that these nodes can be mixed and matched in many different ways. Second, as we showed in the previous section, causal-nodes were combined into different constellations of explanations that changed the function of the of the individual causal-node in the larger reasoning structure**.** In other words, we see this combinability both quantitatively and qualitatively.

Validating Frequency Within People

Finally, we expect there to be some stability in the causal-nodes individuals activate in their pre- and post-responses. What we should address then is, does activating a causal-node in a pre-response *better* predict that a person also activates it in their post-response than we would expect to see if people randomly activated causal-nodes in their responses. For each of the causal-nodes, we calculated the conditional probability that people who activated it in their pre-response also activated it in their post-response and divided by the frequency of that causal-node in post-responses. If this ratio is greater than 1, we see a higher than expected frequency amongst people who also activated the node in their pre-response. As can be seen in the "stability" column, all causal-nodes had a higher than expected degree of stability between pre- and post-responses.

Findings: Association Rule Changes

Now that we have identified students' causal-nodes and hopefully made a convincing argument that these are, indeed, knowledge-pieces of some sort, we can run an Association Rule Mining on our data. The primary output of an ARM are so-called *association-rules*. They take the form of, "if a student activated causal-nodes X and Y, we observed that they *also* activated causal-nodes A and B with a *confidence* of P, a *support* of S, and a *lift* of L". ARM calculates the association-rules for all possible combinations across all responses. Since we have 10 different codes, and each of them can either be present or not, we end up with a total of 2^10 = 1,024 combinations – too many rules to read through in any meaningful way. ARM assists us in navigating this large analytical space by providing three metrics, *confidence*, *support*, and *lift,* that each help identify interesting and important rules: *Confidence* calculates how well the presence of one set of causal-nodes predicts another set of rules, or more formally: the conditional probability that a set of causal-nodes are activated, given the activation of another set of causal-nodes. However, confidence does not consider the aggregate frequencies of causal-nodes, and thus it often overestimates the confidence with which a less frequent causal-node predicts a more frequent one and vice versa.

Rule		Lift			Confidence			Support		
		Pre	Post	Delta	Pre	Post	Delta	Pre	Post	Delta
R1	1->2&6	0.84	1.17	**0.3**	0.37	0.40	0.03	0.17	0.20	0.02
R2	1->0&8	1.21	**1.37**	0.15	0.47	0.40	-0.07	0.22	0.20	-0.02

Table 2. Two Examples of Association Rules (bolded text indicates why they were chosen)

Support and *lift* help us filter out rules in two related ways: *support* simply calculates how often, across the dataset, we see a particular combination of causal-nodes. *Lift* enhances the *confidence* measurement by calculating the ratio between *observed* confidence and the *expected* confidence, given the independent probabilities of each of the two sets of codes. (This is the same approach we used for calculating the within-person stability in the previous section.) Consequently, as we are interested in looking at changes in how students co-activate causal-nodes, the immediately most important measurement is change in *lift* between pre- and post-responses, but looking at lift isolated in pre- or in post- can also help us understand what kinds of constellations of causal-nodes students bring to the unit, and which ones they leave with. We ran an ARM for pre- and post-responses separately, and then calculated the *changes* to *confidence, support* and *lift* between pre- and post-rules. To find interesting and frequent rules, we filtered out rules that had less than .15 support, and included only combinations that we observed in at least 7 out of 41 responses in *both* pre- and post-responses. A full account of all interesting rules is outside the scope of this paper, but in the following, we will give two examples of interesting rules, explain why we think they are interesting and why we chose them, and discuss what they tell us about changes in students' thinking at the classroom level.

R1: Connecting Job Location with Choice and Relative Distances

R1 in Table 2 shows that students who reasoned that jobs are located in cities and not in suburbs were much more likely in the post-response than those who said so in the pre-response, to also say that high income people may live further from workplaces, *and* that high-income people have more choice. We chose this rule because it had the highest change in delta in our ARM. Of course, even when *lift* takes into account the expected frequencies, this change *could* have happened simply because fewer people activated any of the three causal-nodes in their post-responses, and those that did could be the ones spuriously "pulling up" the association rule. However, Table 1 tells us that the frequencies of all three codes are fairly stable from pre- to post- at the classroom level, and in Table 2 we even see a modest increase in support for R1, showing that in absolute numbers, more people co-activated the three causal-nodes in their pre-responses. To us, this indicates that R1 points to a robust change in students' thinking at the classroom level towards connecting the *location of typically higher-paying jobs* with the *choice* and *relative commute distances* of higher income.

Qualitatively, this is a particularly interesting rule to us, because it gets at the part of the question that most people find counter-intuitive: that high-income people often have longer commutes that low-income people. Indeed, one of the reasons we designed the simulation activity was to let students change the infrastructure and the zoning of the city – including the relative position of residential zones and workplaces – to see how common city design patterns (e.g. dense city centres, "green" suburbs far away) lead to this distribution of commute times. While any firm conclusions about causality in this change in thinking would rely on a closer analysis of students' activities during the simulation unit, we speculate that the focus on *placement of zones* in the model helped make this aspect of the phenomenon more salient to students.

Rule 2: Connecting Job Location & And the Desirability of Suburban Life

R2 speaks to the relationship between causal-nodes 1, and 0 and 8. We chose this rule for two reasons: first, R2 has the highest *lift* in post-responses that we identified in any rule, and second, it is interesting because while we see a positive change in lift, we see a drop in confidence between pre- and post-. R2 shows that students who said that jobs are located in cities and not in suburbs were more likely *than expected* to also say that wealthy people live in suburbs *and* that people want to live in places with green areas or good schools. However, as mentioned, we see an absolute drop in confidence, or predictive power, for this rule between pre- and post-. The drop in confidence on its own could mean that more people are activating causal-node 1 in the post-responses without an increase in the co-activation of causal-nodes 0 and 8 in post-, or that fewer people are co-activating causal-nodes 0 and 8 in the post-responses without a corresponding drop in the activation of causal-node 1. Table 1 shows that we do see an increase in the activation of causal-node 1 and a drop in causal node 8 between pre- and post-. However, we see an increase in causal node 0. To us, this is a good example of why looking at lift tells a more nuanced story, and why it is sometimes easier to read association rules backwards when trying to make sense of them: While slightly fewer people co-activated all three causal-nodes in the post-response, the strongest predictor of whether someone did was whether they activated causal-node 1. In other words, this shows a *convergence* across the classroom on the inclusion of causal-node 1 by those who also co-activated causal-nodes 0 and 8.

Qualitatively, this rule is interesting to us, because it shows how students reason not just about the relative position of jobs and residential areas or about high-income people having more choices. It also shows how they perform a kind of meta-reasoning or perspective-taking: reasoning about how *other people reason* about where to live, and how students then activate this reasoning with the rest of causal-nodes. While we do see a strengthening of this particular association rule, we see an overall drop in the activation of causal-node 8. We think perspective-taking is important when reasoning about policy outcomes, and had hoped that our design would have encouraged more of this thinking. We speculate that this might be due to how the underlying logic of the AI agents in the simulation activity was hidden from students, and will in future implementations explore how we can forefront the AI and make this aspect more visible to students.

Discussion, Limitations, and Conclusion

Using Association Rule Mining helped us better identify interesting patterns in changes in students' assembly of causal-nodes into larger explanatory structures, and gave us both some statistics and a vocabulary for measuring and discussing which of these changes were interesting and significant. We found that students' responses seemed converge around particular co-activations, and we speculated how the collaborative simulation activity might have influenced their thinking, and how we could improve on the design. We hope to have provided evidence for our assertion that Association Rule Mining can be a powerful addition to the qualitative researchers' toolbelt when taking a Knowledge Analysis-inspired view of knowledge. In particular, we hope to have shown that ARM can be used in combination with manual qualitative coding to both validate the knowledge-pieces identified at the *level of individual students*, and show changes in the assembly at the *classroom level*.

In contrast to KA's focus on conceptual change at the level of individuals, we only looked at within-student changes when we validated the stability of individual causal-nodes. In future work, we hope to use ARM to first identify important changes at the classroom level, and then use this as a starting point for a more in-depth analysis of the reorganization of knowledge-pieces at the level of individual students. In additional contrast to KA, our knowledge pieces are somewhat less fine-grained, and as we mentioned previously, we took a very permissive approach to coding students' responses. We hope in the future to apply ARM to a both more fine-grained, and more conceptually coherently uniform set of knowledge-pieces.

Because this use of ARM is new, we have limited understanding of exactly how to interpret the stability and change in changes in thinking at the classroom level. Our baseline for comparing changes is always the expected outcome, i.e. the general frequency of a set of causal-nodes in the post-responses. However, we hope to do similar kinds of analyses on more data from this and other reasoning tasks over longer periods of time, and start work on better understanding and measuring the change and stability in knowledge structures, and potentially develop more generally applicable measures.

An important limitation of this study is that students were all attending a private mid-western university, and almost all students reported to have grown up in suburbs. Consequently, the causal-nodes that we identified should be considered expressions of a particular, and fairly limited experience of the world. We hope to expand on this by collecting similar data from other socio-economic or geographic groups.

We do not wish to claim that an ARM can stand on its own as an analysis of student reasoning. But we believe that it provides a tool for measuring and discussing changes in thinking at the classroom level while still anchoring the unit of analysis in a KA-approach to knowledge and thinking that respects the multitude ways in which students can assemble their knowledge. Consequently, we hope that it will find an appropriate place in the computational methods toolbelt of the Learning Sciences.

References

Adams, P. C. (1998). Teaching and learning with SimCity 2000. *The Journal of Geography*, *97*(2), 47.

Agrawal, R., Imieliński, T., & Swami, A. (1993). Mining association rules between sets of items in large databases. In *ACM SIGMOD Record* (Vol. 22, pp. 207–216). ACM.

Bereitschaft, B. (2016). Gods of the City? Reflecting on City Building Games as an Early Introduction to Urban Systems. *Journal of Geography*, *115*(2), 51–60. https://doi.org/10.1080/00221341.2015.1070366

Berland, M., Baker, R. S., & Blikstein, P. (2014). Educational data mining and learning analytics: Applications to constructionist research. *Technology, Knowledge and Learning*, *19*(1–2), 205–220.

Blikstein, P. (2011). Using learning analytics to assess students' behavior in open-ended programming tasks. In *Proceedings of the 1st international conference on learning analytics and knowledge* (pp. 110–116). ACM.

Devisch, O. (2008). Should Planners Start Playing Computer Games? Arguments from SimCity and Second Life. *Planning Theory & Practice*, *9*(2), 209–226. https://doi.org/10.1080/14649350802042231

diSessa, A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, *10*(2–3), 105–225.

diSessa, A. (2002). Why "conceptual ecology" is a good idea. In *Reconsidering conceptual change: Issues in theory and practice* (pp. 28–60). Springer.

diSessa, A., & Sherin, B. (2015). Knowledge Analysis. *Knowledge and Interaction: A Synthetic Agenda for the Learning Sciences*.

diSessa, A., & Sherin, B. (1998). What changes in conceptual change? *International Journal of Science Education*, *20*(10), 1155–1191.

Dorn, D. S. (1989). Simulation Games: One More Tool on the Pedagogical Shelf. *Teaching Sociology*, *17*(1), 1–18.

Frasca, G. (2001). *Videogames of the Oopressed - Videogames as a Means for Critical Thinking and Debate* (Vol. 2001). Master's Thesis, Georgia Institute of Technology.

Frasca, G. (2003). *The Video Game Theory Reader*. Routledge. Retrieved from http://www.amazon.ca/exec/obidos/redirect?tag=citeulike09-20&path=ASIN/0415965799

Gaber, J. (2007). Simulating planning - SimCity as a pedagogical tool. *Journal of Planning Education and Research*, *27*(2), 113–121.

Gee, J P. (2003). What video games have to teach us about learning and literacy. *Computers in Entertainment*, *1*(1), 20.

Gee, J. P. (2007). *Good Video Games and Good Learning* (1st ed.). Peter Lang Publishing.

Hjorth, A, & Wilensky, U. (2014). Re-grow Your City: A NetLogo Curriculum Unit on Regional Development. In *Proceedings of International Conference of the Learning Sciences, ICLS* (Vol. 3, pp. 1553–1554). International Society of the Learning Sciences.

Hjorth, A. & Krist, C. (2016). Unpacking Social Factors in Mechanistic Reasoning (Or, Why a Wealthy Person is Not Exactly Like a Grey Squirrel). Singapore: International Society of the Learning Sciences.

Hjorth, A. & Wilensky, U. (2014). Redesigning Your City-A Constructionist Environment for Urban Planning Education. *Informatics in Education*, *13*(2), 197.

Kim, M., & Shin, J. (2016). The Pedagogical Benefits of *SimCity* in Urban Geography Education. *Journal of Geography*, *115*(2), 39–50. https://doi.org/10.1080/00221341.2015.1061585

Kolson, K. (1996). The Politics of SimCity. *PS: Political Science and Politics*, *1996*(1), 43–46.

Martin, T., & Sherin, B. (2013). Learning Analytics and Computational Techniques for Detecting and Evaluating Patterns in Learning: An Introduction to the Special Issue. *Journal of the Learning Sciences*, *22*(4), 511–520. https://doi.org/10.1080/10508406.2013.840466

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis : an expanded sourcebook*. Thousand Oaks: Sage Publications.

NCSS. (2013). *The college, career, and civic life (C3) framework for social studies state standards: Guidance for enhancing the rigor of K–12 civics, economics, geography, and history*. Author Silver Spring, MD.

NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States*. Washington, DC: The National Academies Press.

Pahl, J. (1991). Finally, a Good Way to Teach City Government! A Review of the Computer Simulation Game “SimCity.” *The Social Studies*, *82*(4), 165–166.

Shaffer, D., Squire, K., Halverson, R., & Gee, J. (2004). Video Games and the Future of Learning. University of Wisconsin-Madison. Retrieved from http://www.discoverproject.net/italy/images/gappspaper1.pdf

Shaffer, D. (2006). *How Computer Games Help Children Learn* (First Edition). Palgrave Macmillan.

Shaffer, D. (2006). Epistemic frames for epistemic games. *Computers & Education*, *46*(3), 223–234.

Sherin, B. (2006). Common sense clarified: The role of intuitive knowledge in physics problem solving. *Journal of Research in Science Teaching*, *43*(6), 535–555.

Sherin, B. (2013). A Computational Study of Commonsense Science: An Exploration in the Automated Analysis of Clinical Interview Data. *Journal of the Learning Sciences*, *22*(4), 600–638. https://doi.org/10.1080/10508406.2013.836654

Sherin, B., Krakowski, M., & Lee, V. R. (2012). Some assembly required: How scientific explanations are constructed during clinical interviews. *Journal of Research in Science Teaching*, *49*(2), 166–198.

Squire, K. (2003). Video Games in Education. *International Journal of Intelligent Simulations and Gaming*, *2003*(2), 1.

Turkle, S. (1997). Seeing Through Computers: Education in a Culture of Simulation. *The American Prospect*, *March-April 1997*.

Wilensky, U. (1999). NetLogo: Center for connected learning and computer-based modeling. *Northwestern University, Evanston, IL*.