Playing a Distributed Epistemic Game: The Forms, Moves, and Tools that Play a Role in Collective Knowledge Construction

Abstract: Discovery is not an individual accomplishment but is built upon preceding and coevolving work. This is true for both singly- and jointly-authored work, though it is more obvious in the work of a group. In this paper we introduce the theoretical framework of a *distributed epistemic game* in which a group of individuals use multiple forms to address a complex problem by looking at the *moves* made during the work on the first Polymath project. By playing this game, scholars are constructing new knowledge in various forms. Describing distributed epistemic games - including the tools, forms, and moves involved - allows us to improve our understanding of knowledge construction as it happens *in situ*.

Introduction

As scientists, we build on the learning and discoveries preceding our own work. Even in single-author publications it is clear that others' work - including prior development of tools, constructs, theories, and methods - plays a key role in discovery. In group work, it is even more obvious that discoveries are a joint activity. In this paper we combine the theoretical frameworks of epistemic forms and games and distributed cognition to build out a new framework of distributed epistemic games using a grounded theory process and use it to describe the moves of players during the first Polymath project.

Complex Problems

A certain set of problems come with an unusual challenge. These problems are not inherently too difficult to be solved, that is to say that a solution is theoretically possible, but have confounded past attempts to solve them because there is not enough space or minds around the problem itself.

In his book, Prelude to Foundation, Isaac Asimov introduces a thought experiment, which we call the cubic meter (m^3) problem (Asimov, 1989, p. 62).

Can you have a relatively small piece of platinum, with handholds affixed, that could not be lifted by the bare, unaided strength of any number of people, no matter how many?

Since “bare, unaided strength” excludes the use of any mechanical assistance (such as levers or gears) there is what seems to be a quick and easy answer. Platinum is heavy enough - with a density of 24,420 kg/m^3 - that the number of people who could fit around the perimeter of the cube would be unable to lift it (ibid, note that other sources list slightly different densities for this element).

Within the book, this thought experiment is used as an analogy for some problems which, because “only so many people could gather round the knowledge,” are doomed to remain unsolved (ibid). It is a thought experiment, but calls out the practical limitation of getting enough minds around a problem. This limit is not absolute, but creates the opportunity for us to solve problems by getting more people to “gather round the knowledge” (ibid). Solutions require more than one mind to work … and more than one person can know. They are similar to the examples presented by Hutchins, where artifacts (such as maps) and teams are necessary because the knowledge required is more than a single mind can observe or know (2006).

In the past attempts to solve these problems have been summarized as “merely difficult, expensive, and unlikely to work” yet possible to solve for any willing to face those barriers (Crichton, 2012, p.73). However, modern tools and functionality allow us to both lower the cost and increase the chance of success by gathering more individuals and their tools. And by increasing the number of individuals, we can pull in a distributed expertise with interdisciplinary experience and increase the likelihood of solving the problem.

In recent years, several such problems have been identified by various political and social groups. These include climate change and pandemic management. These problems affect all of us, and they are big and urgent - ones which need to be addressed in our lifetimes, but they are too heavy for the bare, unaided strength of those who have historically gathered around them. Understanding how a group of individuals can gather around and solve a problem is important as we look to design and streamline the process to address these problems.

Theoretical Foundations

To describe and better understand the behaviors and outcomes in this collective knowledge construction process, we synthesize two existing theoretical frameworks, one characterizing the nature of knowledge-building activities, and the other characterizing the distributed nature of cognition. The synthesis of these two frameworks - epistemic forms & games with distributed cognition - allows us to account for a distributed knowledge construction process.

Epistemic Forms & Games

Collins and Ferguson (1993) presented the framework of epistemic forms and games to describe the way knowledge is built by scholars in both scientific and non-scientific fields. While scholars may not think of their work in this way, the framework of epistemic forms and games suggests that knowledge construction happens as scholars enact moves to play a game which is constrained by a specific epistemic form.

In the same way that the game of Tic-Tac-Toe is constrained by the crosshatch form on which it is played and, within that form, players move by marking an ‘x’ or an ‘o’ in a chosen position, a scholar's knowledge-building game is constrained by a particular epistemic form. The form is a template the scholar "fills out" as they play their game, turning it into an epistemic artifact. Within the form, the available template slots guide the moves they make in playing their game.

An epistemic form might be something as simple as a list of critical elements, as in the case of diSessa's list of p-prims (diSessa, 1993). It might describe the structure of something in space or time, such as a circuit diagram or a stage model (Piaget, 1964). It may describe a process or causal relationships, as would an agent-based computational model (Wilensky & Rand, 2015).

The selection of the form and game is strongly tied to the research question at hand. Based on the knowledge scholars seek to develop, different forms, games, and moves will be most effective.

While we see a connection between our data and the theoretical framework of epistemic forms and games (especially in the visualization of comments as individual moves) there is a lack of explanation for the theoretical impact of playing a game as a distributed or collective process. Research has shown that knowledge construction is frequently distributed (Dunbar, 1997), but the epistemic forms and games framework doesn't account for this distribution. To account for the inherently distributed nature of knowledge-building activities, we turn to the theory of distributed cognition.

Distributed Cognition

Hutchins describes cognition as a distributed phenomenon, one which is not contained within a single biological organism. It depends not only on the individual, but on interactions between that individual and their surroundings. These surroundings include other individuals, tools/artifacts, and the environment itself. Each of these takes some part of the cognitive load. Hutchins (2006) gives an example of how even an ant can appear more intelligent than they would be as an individual because of interactions with "the residua of the history of its ancestor's actions" (p.169). The residua (or knowledge) of past interactions changes the environment in such a way as to make all future behavior seem more intelligent.

This knowledge persists in adjustments to artifacts, environment, and behavior. With these changes, future thinking in the environment is also changed. This creates repositories of knowledge not only in the individual mind, but in group behaviors, changes in artifacts, and differences in the environment itself. Artifacts distribute cognitive work over time and multiple individuals - including initial creators and those who adjust or use it moving forward. These artifacts provide a way for information to travel across groups and time (Baltzersen, 2022). Future individuals have a changed environment - with different artifacts and behavior patterns - in which to be thinking.

Distributed Epistemic Forms & Games

In the present work, we discuss the idea of distributed epistemic forms and games, which looks at epistemic games through the lens of distributed cognition. While the forms themselves are not distributed, they are used as epistemic games played collectively by members of a group differing in expertise, location, and time/duration of interaction. Because these games are not bound by any of the constraints discussed above, they allow for a larger number of participants and interactions.

This framework can be used to model distributed epistemic games, identifying, for example, the moves of which they are composed and the organization of the moves across players, time, and space. Within this paper, we discuss the idea of a tactic: a collection of moves made in an epistemic game. In chess, masters will play multiple tactics, each of which is intended to add a threat to the opponent and advance the player toward victory. In distributed epistemic games, the intent is not to threaten an opponent, but to take multiple paths towards the goal of the larger game. While some of these tactics are deliberately planned, others emerge as a result of the interactions between individuals. The emergent tactics are of special interest specifically because they are not planned, but rather are an organic (if complex) property of a distributed epistemic game.

The Problem: DHJ and the first Polymath

The DHJ Problem

For an overview of the density Hales-Jewett (DHJ) problem we turn to Reinventing discovery by Michael Nielsen (2014). In the appendix, he gives an explanation of the problem, including some visuals and context for what made the solution found by the Polymath project a valuable addition to mathematical understanding. We have recreated the visuals here along with an explanation of the problem and a mention of the pre-existing proof.

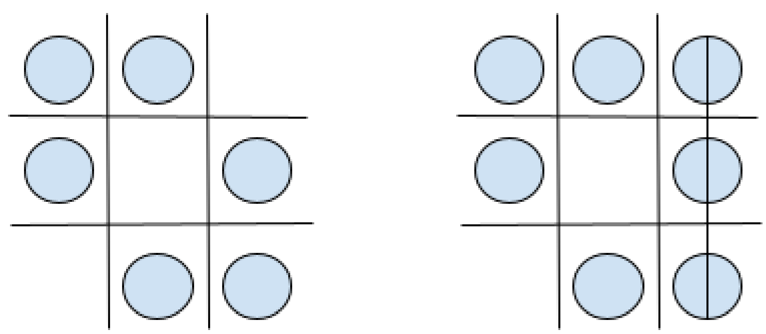

Figure 1. *Two example two-dimensional, three-by-three grids: one with six spots filled, one with seven showing a line that goes through three of the seven dots. A three-by-three grid may have as many as six spots filled before at least one line can be drawn through three of the dots.*

Starting with a two-dimension three-by-three grid (Figure 1, left) we can create a “line-free” arrangement of dots as long as we fill six or fewer of the nine available places. Once we fill the seventh place (Figure 1, right) there will be at least one possible line we can draw through three dots. As we add dimensions, for example going from a two-dimensional grid to a three-dimensional grid, the line-free configuration (the largest number of filled spots possible before you can draw a line through at least three of the dots) becomes a smaller and smaller proportion of the overall size. In fact, as the n number of dimensions gets larger, the largest possible line-free configuration becomes smaller and trends towards zero. Before the Polymath project, there was an existing proof using an obscure branch of mathematics (ergodic theory) and the objective of the Polymath project was to find an alternate, and hopefully more accessible, proof.

A Case Study of Collective Knowledge Construction - The First Polymath Project

Imagine a group of individuals, spread across continents, working together for the fun of it as they seek to solve a problem (as yet unsolved) but also to discover whether it is possible to work together in such a distributed way and accomplish anything meaningful. At the outset, one member wonders aloud whether it can be done - and declares that the project would be successful (even if the problem were unsolved at the end as well) if they found a way to work together. ("Is Massively Collaborative Mathematics Possible?," 2009)

The first Polymath project was proposed as a quest not necessarily to solve a problem, but to see whether a group of individuals could "engage in massive and collaborative problem solving" ("Is Massively Collaborative Mathematics Possible?," 2009). The chosen problem was to create a proof for the density Hales-Jewett problem (Polymath, 2012). By agreement, the work was mostly restricted to blog posts and comments, though a few instances exist in which individual contributors left to do some computations or to update a wiki dedicated to the project and came back with a report.

At a very early stage of the project, individuals started numbering key comments and separating into named threads. This served as a sort of ongoing rank-and-sort process to identify those comments which made key contributions and to keep different approaches contained. It is important to note, however, that this numbering process had exceptions, for example when users would deliberately not number a comment to avoid "filling up" the thread before they were ready since the threads had a user-imposed limit of 100 numbered comments. Comment numbers and timestamps/names were used to reference other comments in a reply or continuation of the thought process.

Though discussions around crediting the work (especially the problem of judging contributions for the purposes of researcher reputation & job-seeking) happen throughout the process, the individual contributions can be traced and followed by reviewing the posts and comments on the two blog sites which hosted the project.

Over a decade later, this project has become an academic urban legend - an example of the remarkable possibilities of collaborative work. Books and papers (Malone, 2018; Nielsen, 2014; Cranshaw & Kittur, 2011) will point to this as a case study in collective intelligence, where the result of the group was better than what any member could have achieved individually … or even the sum of individual achievements. This project will be mentioned in the same discussions as others such as FoldIt and the DARPA Red Balloon challenge (Koepnick et al., 2019; Madnick, 2022). And, like those projects, researchers will claim that something in the collective or the distributed feature made it successful.

There are some key differences about this project, however, which draw our interest. First, it was not building on the success of the platform, format, or group. There was a learning curve because none of these aspects had been used in this way before. Second, there was a large and documentable range of expertise among participants with several having no published work before engaging with the Polymath project, while a handful had hundreds. (Cranshaw & Kittur, 2011). Finally, because of the repositories of blog sites still in operation, it is possible to pull all data for individual posts and comments. Since the work on the project happened exclusively within the structure of two blog sites, there is public access to what is effectively the transcript of the entire project. This gives a unique opportunity to study the interactions and process - to better understand what made this work - and to work with theory to generate meaning based on empirical data.

Methodology Review

For the reader's convenience, we will give a brief overview of methods here. More curious readers are directed to other papers in which there are more thorough treatments of the various methods used (see, for example: Matthews, Nguyen, & Swanson, 2023; Matthews & Swanson, 2023; and Matthews, in preparation).

Data (blog posts & their associated comments) were initially extracted from the two blog sites (https://gowers.wordpress.com/; https://terrytao.wordpress.com/) manually. After reviewing the posts, a python script was created to automate the extraction and some basic transformations such as date formatting and text cleaning.

Once the data were extracted, we used a grounded theory building process to identify meaningful patterns, actions, and outcomes (Glaser & Strauss, 2010; Haig, 1995; Strauss, 1987). Specifically we focused on an iterative and recursive coding which allowed for the refinement of codes over time. While this grounded theory process was the primary method, we also drew some methodological inspiration from Knowledge Analysis (diSessa, Sherin & Levin, 2016) - especially the distinction between *microanalytic* work, or describing key moments, and *microgenetic* analysis describing change over time (*ibid*). Even though *knowledge analysis* historically focuses on the individual learner, we found the distinction helpful in our grounded theory building process. In this paper we focus on what would be described as *microanalytic* work, with *microgenetic* analysis as a goal for future research efforts.

Sans Coding

Using these methodologies, and the data for the first Polymath project, we start at the level of the full game - pulling all available blog post & comment data for the project.

Our next step is to evaluate the data through several passes of coding. This methodology, which we call sans coding, is described in more detail in another paper (Matthews, in preparation) so will only be discussed at a high level with example figures here.

416. Obstructions to uniformity., It also occurs to me that Boris's thought processes could be extremely helpful here. Did that example just jump out of thin air? If not, then were there some preliminary "philosophical thoughts" that led to it? If the

Figure 2: *In the first pass of sans coding we highlight parts of the posts that stand out to draw attention for future analysis.*

trimmed_move	post_text
suggested reading	145. Sperner., Hi Tim, re your #142. It is known that for all , among sets and on the sphere of fixed measures, the ones which minimize the measure of pairs with are the "concentric" caps. Note that in high dimensions, having inner product is very much
clarification	146 Sperner, Terry @118 142, Yes, We need more structure to apply Endre's trick. It was easy to apply when (almost) every element is a center of symmetry in the following way; Choose an element of your set (center). Any other element has a unique
considering potential paths	147. Sperner, Terry — indeed. In fact, I came to a similar realization in the DHJ context at the end of the first paragraph of 140. But I suppose one could regard things in another way and say that we're trying to get the density to tend to zero (very very
einstein approximation	148. Sperner, As I was sitting through my daughter's piano lesson, my mind wandered back to pair removal in , and I found myself wondering whether we were already done. Almost certainly not, but in the spirit of keeping thoughts public, let me attempt to
coalesce	148. Obstructions to uniformity, It took me a while to catch up with this thread – this is perhaps one thread which is particularly ripe for a fresh start with a post summarising previous progress and listing the known examples., Here are the examples of
clarification	149. Obstructions to uniformity., A couple of things I didn't quite understand there Terry. If you pull over the obstructions from (as suggested in 116), do you call that a pseudorandom deviation? Where does that example fit into your scheme above? And
suggested reading	150. Re the last sentence in Tim's #142: I mentioned this in #82 but perhaps I should have pointed to a later-on paper in the Irit Dinur oeuvre:, Click to access intersecting.pdf, Here she and Friedgut give a mostly complete characterization of large
review	151. Sperner, Ryan, I didn't manage to find an obviously relevant part of the paper you linked to in 82, but the reference you've just given does indeed look promising. Do you know enough about their argument to know whether it is robust? That is, if

Figure 3: *The second pass of sans coding focuses on coding every interaction without simultaneously editing the codes.*

Category	Code	Note	Example
check understanding	asking questions		309. Pair removal in Kneser graphs., (
check understanding	answer or statement	(UNLESS answer has an 'aha' or 'huh' moment)	327. Kneser, Yes, so if you can prove t
check understanding	both question and answer	both in the same post	589. Dimensions, I agree with Terry, ir
check understanding	correct prior answer		Oops, on second reading it's Pan
working the problem	computer-proof	incl. reporting code run	953., My program is still running for tl
working the problem	thread management		300. One of the main reasons for this
working the problem	request or response		Could someone post the 387 Pare
working the problem	human-proof	reporting work done elsewhere by hand	And congratulations to Kristal for

Figure 4: *For the third pass, we group and edit the codes from the second pass to see overall patterns. This process is where we create a code table, complete with example moves for every code.*

move_category	move_code	post_text	
meta	justify problem	Of course, one might say, there ar	
meta	who gets credit	Two quick questions., 1. How do c 1. threaded posts, which are of co	
check understanding	answer or statement	Some answers off the top of my h	

Figure 5: *Finally, the fourth pass involves coding every interaction with the grouped & edited codes from the code table.*

Once this is complete, we have standardized the coding for the data as a whole, and the codes can be used to identify and characterize tools, forms, players, moves, and tactics within an example of a distributed epistemic game.

Parts of the Game: Tools, Forms, Players, Moves, and Tactics

The Role(s) of Tools

Tools play several key roles in a DEG. Some tools (such as blog sites) are distributing, meaning they facilitate distributed cognitive work by providing a common site on which a variety of individuals can interact. Other tools (such as software or hardware) are discreet, meaning they assist and improve the process without necessarily providing a common site for individuals to interact. In this section we will review several roles, and then discuss specific tools which are part of the first Polymath DEG.

Tools as Cognitive Prosthetics: Increase the Reach

Each new set of tools presents new opportunities for expanding the work that is possible for players of a DEG. There are many ways that tools facilitate this expansion, including by acting as a repository of knowledge (since any adjustments/refinements to the tool represent increased knowledge of the users), acting as a transmitter of knowledge , serving as an object to think with, and providing the structure that allows for a distribution in time and space.

In a cognitive ecosystem, both organic and inorganic elements play a role in the thinking that can/does happen. Tools can be broadly defined as “non-living matter [used] to form cognitive prosthetics” (Garfield, 2022). These prosthetics can then be used to increase our sensitivity to external stimuli (as in the case of a spider’s web or a radar array), to offload cognitive tasks like memory (as in the case of navigation charts or virtual assistants such as Siri and Alexa) and computation (as in the case of slide rules, calculators, and computers), and to transfer knowledge (as in the case of writings and the changes in the tool itself). While none of these examples are living matter, they all play a key role in making sense of the environment through receiving input, computing change, and storing knowledge.

Tools for Memory Extension: Increase the Storage

Vygotsky - well ahead of our usage of tools such as digital assistants - describes memory-extension behavior as "the process of memorizing by forcing an external object to remind her of something" (1981, p.57). This is an example of a "basic characteristic of human behavior in general is that humans personally influence their relations with the environment and through that environment personally change their behavior subjugating it to their control" (ibid). By using a digital tool in this way, learning (or a change in behavior) happens directly and personally through human interactions with the environment.

Tools for Computation: Increase the Speed & Accuracy

One of the ways tools can play a role in the sense-making process is through the existence of constraints. Whether physical or digital, all tools have limitations which constrain how/where/when they can be used. These constraints can be frustrating, or they can be used to create the sort of catalyst that precedes realization. An example of this happens in the Private Universe work (Schneps et al., 1989) where a middle school student “pulls out the models” (see Figure 6 below) and that becomes a catalyst for understanding.

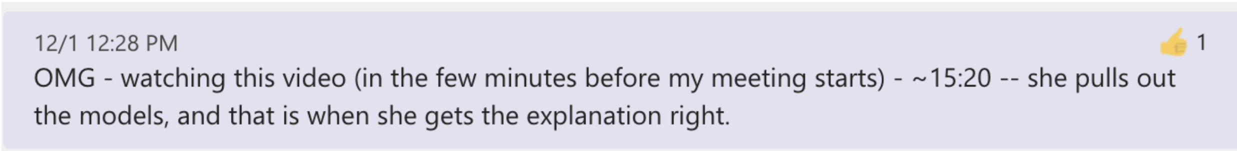
12/1 12:28 PM 1
OMG - watching this video (in the few minutes before my meeting starts) - ~15:20 -- she pulls out the models, and that is when she gets the explanation right.

Figure 6: *Personal Correspondence, author, December 2020 - reaction to the student using models to explain the phenomenon*

Constraints inherent in the tools (such as cohesion and physical existence) act as a sort of rigid landmark (Arnell, Swanson, & Edwards, 2022) which forces cognitive change to happen. Instructors can make use of these constraints by using balls to teach axial tilt, and legos to teach chemical reactions.

Tools for Translation: Increase the Dimensions

Expanding to collective learning (change in intelligence) from the individual learning perspective, tools serve an additional key purpose of translation by acting as a common substrate to facilitate the transfer of knowledge or information across time and fields.

In Mind & Society, Vygotsky describes the evolution of pointing from a failed attempt to reach a goal oneself to a signal intended to mobilize other humans for assistance (p.62). In this evolution, the goal (an inanimate object in his example) demonstrates the line between the ZUA and the ZPD - the point at which additional assistance is needed to reach the objective. Between the ZPD and the ZCP, an object serves not as a boundary but as a way to transfer and translate the thinking process, though it remains in the role of facilitating " actual relations between human individuals." (p.62). It is the "objectifications of human needs and intentions already invested with cognitive and affective content" (Wartofsky, 1973 as cited in Cole, 1996)

The tool may be something digital (such as a blog post) or a set of "tangible bits" which can interact with "radical atoms" in the creation of cognitive change (Hines et al., 2011 as cited in Malone, 2018). It may be as simple as a physical chart, or as abstract as a word or term. In any form, the tool is serving a cognitive function of increasing capacity and catalyzing change by the challenge of human memory not being sufficient to contain human knowledge (Baltzersen, 2022) - of the total knowledge exceeding "what could be known by any individual" (Hutchins, 2010, p.111).

Creating a tool allows for the knowledge to be materialized in a way that it can be stored, adjusted, spread, and reused (ibid). Over time, these adjustments take the form of continuous correction or modification - storing the knowledge of generations of users within the tool itself. From physical objects to scientific knowledge, tools are continually up for evolution and development (Horner, 2010).

Additionally, the tools are able to spread cognitive effort across individuals and across time, as in the example from Hutchins "The cartographer has already done part of the computation for every navigator who uses his chart." (2005, p.173). One of the ways that epistemic games are distributed is by spreading across generations of time in the tools that have been created and continue to be refined by human users. A tool may be new or ancient, but when we use it we are connecting to the wisdom of individuals before us and spreading cognitive load backwards across history.

The Tools - Both *distributing* and *discreet/discrete*

Within this specific *distributed epistemic game* there is a collection of tools which have a key role. Some of these tools are *distributing* in the sense that they facilitate the *spreading* of knowledge and information to the group. Others are both *discreet* in the sense that they are not *spreading* information and *discrete* in the sense that they exist in a separate and distinct way. Several of these tools (both *distributing* and *discreet/discrete*) are listed with some of their important characteristics below:

- Blog sites (*distributing*) - the constraints of post/comment format, lack or presence of LaTeX support, and ability to track comments across various threads
- Wiki sites (*distributing*) - ability to store/edit more permanent copies of work; options to constrain participation to reduce spam
- Maple and other mathematical software (*discreet/discrete*) - allowed for computation at scale (quickly became important as the game continued); gave an alternate path-to-proof so that finished results could be checked both by hand and by computer
- Physical computing machines (*discreet/discrete*) - needed to run the scaled computations

The Forms

Epistemic Forms are the things that scholars are trying to fill out or complete or create in the knowledge construction process. A form is not - in itself - 'distributed', though as an abstract artifact it does *support* the distribution of cognitive work through embedded knowledge of information organization. It becomes part of a *distributed epistemic game* when multiple forms are in play - with multiple players - simultaneously and **changes/moves in one form have a cascading or rippling effect across other forms**.

Within this specific game, the following forms have been identified:

- Lookup Table - similar to the list form discussed in Collins & Ferguson (1993); the lookup table is a specialized version of the list form where the list values are constrained to a specific type and format. The lookup table is created during an epistemic game, and may then be used as a tool or reference for players moving forward.
- Computer Code - a script or program which is created or adjusted during the process of playing the game.
- Mathematical Proof - a tightly constrained, and field-specific, form which represents a final target for the work of the distributed game

While the tools and forms in a DEG might not differ from those of a traditional epistemic game, the players, moves, and tactics have more distributed sources and effects.

The Players

The Polymath project saw a wide range of participation volume from the players. Previous researchers have reviewed the data at a higher level and found a total of 39 users (29 of whom could be identified) who made at least one numbered contribution to the process (Cranshaw & Kittur, 2011).[1] Among these there is a large spread in expertise - measured by Cranshaw & Kittur as the count of refereed publications. Numbered contributions came from Fields medalists and hobbyists alike.

The most simple form of a DEG is one in which multiple individuals are 'playing the game' together, bringing different perspectives, expertise, and moves together to reach a common goal.

Moves & Tactics

Moves

The moves in a DEG "can act as a mediating device that controls the pattern of availability of data for the other" (Hutchins, 2006, p.340). Each time a player makes a move, it has the potential not only to move the larger game forward, but to change the forms/moves available for other players and future moves.

Below is a partial list (with examples) of moves which show up in the data for the first Polymath project.

- *Suggested readings* - in this move, a player suggests or responds to readings (usually external to the DEG) to further clarify a position or method.

  602., Dear Terry,, I think it's a very interesting idea to have an online reading seminar. Of course, it requires us to do the reading, which is quite a big investment to make, and presumably why you envisage a sedate affair. I think what would be most useful to me is basically what you encourage: rather than reading the paper very carefully and then explaining bits I understand here, I would prefer to try to use this post and its associated comments to read the paper sloppily and present my very partial understanding, or even complete misunderstanding, in the hope that others will be able to say, "Yes, you're getting there — what you need to grasp next is this," or "No, that's not quite it — if that were the case then so would this be but it clearly isn't," and things like that.

- *Updating knowledge artifacts* - in this move, a player updates an artifact being used to capture changes in knowledge throughout the game.
  260. Updates to spreadsheet, I've updated the spreadsheet to take into account the above developments (in particular, adding entries for upper and lower bounds

- *Thought Process* - in this move, a player reflects on and explains their own thought process in earlier work.
  418. Answer to 416: My thought process was simple: First the analogue of Tim's conjecture for Sperner is false (if one thinks of Sperner as DHJ(2,1), then it is not true that obstructions come from DHJ(2,0) because there is no DHJ(2,0)), and hence the conjecture should be false.

[1] The numbering of comments started early in the project and numbers were self-assigned by the comment author on comments which made meaningful progress towards the goal. It should be noted that sometimes authors would deliberately not number comments to avoid hitting the 100-comment limit on numbered posts.

Tactics

While many moves happen as stand-alone events, it is possible for a collection of moves to form a *tactic*. Moves (as stand-alone events) may not cause or involve much interaction between the parts of the DEG, but tactics engender interaction. There are three tactics (one planned or intentional, two emergent) we discuss within the setting of a distributed epistemic game. The first is called *ask and respond* and is a very basic check for understanding. The second, which may emerge from ask and respond is called the *pivot*, and is described in more detail in other papers but will be mentioned here (Matthews & Swanson, 2023; Matthews, Nguyen, and Swanson, 2023). The final tactic we discuss is the *coalescence* which is a sort of cognitive collision where ideas are transferred within/between conversations and - most excitingly - across fields. This lays the groundwork for meaningful innovation to occur at the edge of a field - or where it collides with others (Painter et al., 2021).

- Planned
    - *Ask & Respond* - The ask and respond tactic is one of the most basic in a distributed epistemic game. This tactic would normally involve at least two individuals and happen *with minimal disruption* to the process of playing the epistemic game. The basic objective of this tactic is to confirm that the learner (in this case the collective) is still on the correct pathway and making correct assumptions, similar to the use of instructional landmarks discussed by Arnell (Arnell, Swanson, & Edwards, 2022). In individual learning activities these are usually formatted differently, often using the initiate-respond-evaluate framework where the learner is asked to make sure they understand, rather than the learner asking the question. These knowledge checks are usually quick and minimally disruptive to the learning process and have been demonstrated to have significant and causal influence on both immediate learning (Koedinger et al., 2016, 2018) and longer term retention (Van Campenhout et al., 2021).
        - *Ask* - one player checks their understanding or asks a clarifying question
          15. Regarding lower bounds, fixing the numbers … gives better lower bounds, right?
        - *Respond* - another player responds
          16. Hmm, Gil, you're right; so we get …

- Emergent
    - *Pivot* - The *pivot* tactic is described extensively in earlier work (Matthews & Swanson, 2023; Matthews, Nguyen, and Swanson, 2023) so we will mention it only in passing here. A pivot may begin looking very like an ask and respond, but the *key difference is that a pivot disrupts the thinking process*. This disruption - a change in direction, interaction volume, or other characteristics - emerges through interactions between individuals and is the defining trait of a *pivot*.
        - *Ask* - the *pivot* starts with a question, similar to the *ask & respond*
          309. Pair removal in Kneser graphs., Can someone help me with this dumb question?
        - *Interesting* - the difference is that the *response* in the *pivot* disrupts the thinking process, in this case stopping a monologue and changing it to a discussion

> 310. Pair removal in Kneser graphs., … all I can say is that I can't see a flaw in your reasoning. If it's correct, then it shows something rather interesting.

- *Coalesce* - The final tactic we discuss in this paper is the *coalescence*. This happens when individual players connect across posts, forms/threads, and fields - such as when players bring techniques from genetics to bear on problems in mathematics. This creates a *cognitive collision* as the ideas, assumptions, and patterns collide across different individuals, forms, and fields. These collisions enable greater innovation, as shown in Painter et al., (2021) because they bring together different ways of looking at the problem based on different training and perspective. The result of combining all of these disparate parts is the creation of knowledge which wasn't there before - which can be traced back to each of its precursors and (by existing) connects ideas and fields which were not connected before (Borges, 2000). In this way, it works as a kind of bidirectional translation across disciplines, especially one in which the focus is not on individual learning (Heath, 1983).

> 716. Genetic algorithms … it would be amusing to see how well a genetic algorithm would fare with trying to maximise the size of a line-free set. It's tempting to use blocks … as "genes" that can be swapped around between competitor sets …

Conclusion

In this paper, we have used the new framework of a *distributed epistemic game* or DEG in a grounded theory analysis of the first Polymath project. We have described *tools, forms, players, moves,* and *tactics* which play a role in the DEG. At this point we suggest several options for future analysis using the DEG framework.

Our current analysis is close to the framework of *microanalytic* work in Knowledge Analysis (KA) where we are describing moments during the learning process, but we are looking at the learning of the collective, rather than the individual (diSessa, Sherin & Levin, 2016). Future work would do well to move to a *microgenetic* style of analysis where the change over time is plotted and described.

For our analysis here, we have looked at each *post* as a single *move* in the game - describing the post itself and the ways in which a single *move* can combine with others into a *tactic*. Several posts, however, can also be seen as a collection of moves, which provides an opportunity for future analysis to reduce the grain size and code many moves within a single post.

Our final recommendation for future analysis involves adding a coding scheme for the *vector* of individual *moves* (at any grain) so that we can further describe *moves* by showing the progress they make (both direction and speed). This particular analysis could also be applied to *tactics* as well as the individual *moves*. We expect that this particular analysis will require methods for collaborating with experts across disciplines to verify the evaluation. At another grain, future work in this collaboration will be its own *distributed epistemic game*.

References

Adams, S. (2004). God's debris: A thought experiment. Andrews McMeel Pub.

Arnell, J., Swanson, H., & Edwards, B. (2022). Instructional Landmarks: Describing a Novel Intermediate Knowledge Structure for Physics Learners. In Proceedings of the 16th International Conference of the Learning Sciences-ICLS 2022, pp. 1393-1396. International Society of the Learning Sciences.

Asimov, I. (1989). *Prelude to Foundation*. Random House Worlds.

Baltzersen, R. K. (2022). *Cultural-historical perspectives on collective intelligence: Patterns in problem solving and innovation*. Cambridge University Press.

Borges, J. L. (2000). Labyrinths: Selected stories & other writings. New Directions.

Crichton, M. (2012). Jurassic Park (20th Anniversary). Random House.

Cole, M. (1996). Cultural psychology: A once and future discipline. Belknap Press of Harvard University Press.

Collins, A., & Ferguson, W. (1993). Epistemic forms and epistemic games: Structures and strategies to guide inquiry. Educational psychologist, 28(1), 25-42.

Cranshaw, J., & Kittur, A. (2011). The polymath project: Lessons from a successful online collaboration in mathematics. 1865–1874.

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

DiSessa, A., Sherin, B., & Levin, M. (2016). Knowledge Analysis: An Introduction. DiSessa, A. A., Levin, M., & Brown, N. J. S. (Eds.). (2016). Knowledge and interaction: A synthetic agenda for the learning sciences. (pp. 30–71). Routledge.

Dunbar, K. (1997). How scientists think: On-line creativity and conceptual change in science.

Frank A, Grinspoon D, Walker S (2022). Intelligence as a planetary scale process. International Journal of Astrobiology 21, 47–61. https://doi.org/10.1017/S147355042100029X

Garfield, M. (Host). (2022, August 18). Caleb Scharf on The Ascent of Information: Life in The Human Dataome [Audio podcast]. Complexity. https://santafe.edu/culture/podcasts#Complexity

Glaser, B. G., & Strauss, A. L. (2010). The discovery of grounded theory: Strategies for qualitative research (5. paperback print). Aldine Transaction.

Haig, B. D. (1995). Grounded theory as scientific method. Journal of Philosophy of Education, 28(1), 1–11.

Heath, S. B. (1983). Ways with words: Language, life, and work in communities and classrooms. Cambridge University Press.

Hines, J., Malone, T., Gonçalves, P., Herman, G., Quimby, J., Murphy-Hoye, M., Rice, J., Patten, J., & Ishii, H. (2011). Construction by replacement: A new approach to simulation modeling: J. Hines et al.: Construction by Replacement. System Dynamics Review, 27(1), 64–90. https://doi.org/10.1002/sdr.437

Hutchins, E. (2006). Cognition in the Wild. MIT press.

Hutchins, E. (2008). The role of cultural practices in the emergence of modern human intelligence. Philosophical Transactions of the Royal Society B: Biological Sciences, 363(1499), 2011–2019. https://doi.org/10.1098/rstb.2008.0003

Hutchins, E., & Klausen, T. (1996). Distributed cognition in an airline cockpit. In Y. Engeström & D. Middleton (Eds.), Cognition and Communication at Work (1st ed., pp. 15–34). Cambridge University Press. https://doi.org/10.1017/CBO9781139174077.002

Piaget, J. (1964). The Early Growth of Logic in the Child: Classification and Seriation (B. Inhelder, Ed.) (1st ed.). Routledge. https://doi.org/10.4324/9781315009667

Is massively collaborative mathematics possible? (2009, January 27). Gowers's Weblog. https://gowers.wordpress.com/2009/01/27/is-massively-collaborative-mathematics-possible/

Koedinger, K. R., McLaughlin, E. A., Jia, J. Z., & Bier, N. L. (2016). Is the doer effect a causal relationship? How can we tell and why it's important. 388–397.

Koedinger, K. R., Scheines, R., & Schaldenbrand, P. (2018). Is the Doer Effect Robust across Multiple Data Sets?. International Educational Data Mining Society.

Koepnick, B., Flatten, J., Husain, T., Ford, A., Silva, D.-A., Bick, M. J., Bauer, A., Liu, G., Ishida, Y., Boykov, A., Estep, R. D., Kleinfelter, S., Nørgård-Solano, T., Wei, L., Players, F., Montelione, G.

T., DiMaio, F., Popović, Z., Khatib, F., … Baker, D. (2019). De novo protein design by citizen scientists. Nature, 570(7761), 390–394. https://doi.org/10.1038/s41586-019-1274-4
Madnick, S. (2022). Bursting a few balloons regarding the famous DARPA red balloon challenge. Communications of the ACM, 65(3), 33–34. https://doi.org/10.1145/3517127
Malone, T. W. (2018). Superminds: The surprising power of people and computers thinking together (First edition). Little, Brown and Company.
Malone, T. W., Laubacher, R., & Dellarocas, C. (2009). Harnessing Crowds: Mapping the Genome of Collective Intelligence (SSRN Scholarly Paper No. 1381502). Social Science Research Network. https://doi.org/10.2139/ssrn.1381502
Matthews, Jenna A., Nguyen, Ha, & Swanson, Hillary. (2023). Uncovering Features of Discourse that Increase Interactions. Companion Proceedings 13th International Conference on Learning Analytics & Knowledge (LAK23). Learning Analytics & Knowledge (LAK23). https://www.academia.edu/96048886/Uncovering_Features_of_Discourse_that_Increase_Interactions
Matthews, Jenna A. & Swanson, Hillary. (2023). The Pivot: Identifying Emergent Tactics in Distributed Epistemic Games. Computer-Supported Collaborative Learning, Montreal. https://www.academia.edu/98909334/The_Pivot_Identifying_Emergent_Tactics_in_Distributed_Epistemic_Games
Matthews, Jenna A. (in preparation). Sans Coding using a Distributed Epistemic Game (DEG)
Nielsen, M. A. (2014). Reinventing discovery: The new era of networked science (First paperback printing). Princeton University Press.
Painter, D. T., Daniels, B. C., & Laubichler, M. D. (2021). Innovations are disproportionately likely in the periphery of a scientific network. Theory in Biosciences, 140(4), 391–399. https://doi.org/10.1007/s12064-021-00359-1
Parnafes, O., & diSessa, A. A. (2013). Microgenetic Learning Analysis: A Methodology for Studying Knowledge in Transition. Human Development, 56(1), 5–37. https://doi.org/10.1159/000342945
Parslow, G. R. (2013). Commentary: Crowdsourcing, foldit, and scientific discovery games: Crowdsourcing, Foldit, and Scientific Discovery Games. Biochemistry and Molecular Biology Education, 41(2), 116–117. https://doi.org/10.1002/bmb.20686
Polymath, D. (2012). A new proof of the density Hales-Jewett theorem. Annals of Mathematics, 1283–1327.
Rutherford, A., Cebrian, M., Hong, I., & Rahwan, I. (2020). Impossible by Conventional Means: Ten Years on from the DARPA Red Balloon Challenge. https://doi.org/10.48550/ARXIV.2008.05940
Schneps, M. H., Sadler, P. M., Woll, S., & Crouse, L. (1989). A private universe. Astronomical Society of the Pacific.
Strauss, A. L. (1987). Qualitative analysis for social scientists. Cambridge University Press.
Van Campenhout, R., Johnson, B., & Olsen, J.. (2021). The Doer Effect: Replicating Findings that Doing Causes Learning. Proceedings of the Thirteenth International Conference on Mobile, Hybrid, and On-line Learning, eLmL 2021, pp.1-6, 2021, Nice, France. ISBN: 978-1-61208-873-0;
Vygotskij, L. S., & Cole, M. (1981). Mind in society: The development of higher psychological processes (Nachdr.). Harvard Univ. Press.
Wartofsky, M. (1973). Models. Dordrecht: D. Reidel
Wilensky, U., & Rand, W. (2015). An introduction to agent-based modeling: Modeling natural, social, and engineered complex systems with NetLogo. The MIT Press.