Universal Tools for Measuring Games and Learning: Dynamic Causal Nets

Dominicus Tornqvist  
School of Information and Communication Technology  
Griffith University  
Brisbane, Australia  
dominicus.tornqvist@griffithuni.edu.au

Lian Wen  
School of Information and Communication Technology  
Griffith University  
Brisbane, Australia  
l.wen@griffith.edu.au

Jennifer Tichon  
School of Education and Professional Studies  
Griffith University  
Brisbane, Australia  
j.tichon@griffithuni.edu.au

Abstract—How can one measure the learning outcome of playing a serious game? We need an objective measure of the learning contents included in a game. The research is diverse, utilizing vastly different games to teach various different kinds of knowledge and skills. This makes it difficult to compare and generalize studies, lacking any established formal tool of analysis. This problem requires the design of an abstract and objective measurement of the quantity of learning material independent of the learning domain. Based on cognitive research on causual Bayes nets (CBNs), this paper proposes using dynamic causal nets (DCNs) to model an abstract knowledge base, which could be mapped to many different domains of learning. We also apply Kolmogorov Complexity (KC) as an approach to measure the content of the abstract knowledge base. This work will establish a theoretical foundation for future research of serious games.

Keywords—learning; education; serious games; causal bayes nets; modeling; complexity; dynamic causal nets; Kolmogorov complexity; knowledge base, knowledge representation

I. INTRODUCTION

In health, education and training, serious games have been touted as a promising tool to teach behaviors and knowledge, such as mental coping strategies, first aid, nutrition and healthy lifestyle information. Different studies may investigate aspects of motivation, player behaviors, or learning. For example, games to promote exercise, rehabilitation, hygiene, or coping strategies may be more interested in cultivating certain behaviors in players. But by far one of the most central concerns of serious game design (and game design generally) is that of learning. After all, if the player doesn’t know how the game works or what to do, they cannot play it in the first place. And educational games might try to teach about nutrition, first aid, or general information about a newly-diagnosed chronic condition. Due to its centrality to serious game design, learning was the focal example of this study. But formal models of games could be used as a common language to investigate any number of aspects of games.

Definitely games are changing how we work, play, socialize and learn, but are they teaching what we propose they are? Commentators acknowledge that the way people learn in the spaces of online games is fundamentally different than what is considered as standard pedagogical practice [1], yet the push of these games into the educational sector keeps expanding. It is proposed the educational power of the games lies inherently within their ability to teach player’s the skills of using their imagination [1]. Schools have been using Minecraft for several years while in parallel adult education has been using serious games for work place learning [2].

Yet despite their widespread use, knowledge underpinned by objective research is limited or non-existent on those game mechanics that result in increasing knowledge. Those developing games for education or training should be devising associated evaluations [3] to ensure expected learning outcomes occur. Currently, it is difficult to interpret the diverse and somewhat mixed evidence for game-based learning (for reviews, see [4], [5], [6], [7], [8]). It is made yet more difficult by the absence of any standardized methods of formally modelling the systems, especially the learning contents, in such studies. There is a general absence of any unifying underlying theory to tie these studies together, to make them in some way comparable to each other using objective measures of the properties of these games or the learning involved.

Fortunately, the literature on systems theory can provide a good starting point. What is needed is a universal modelling method that can be applied broadly, to virtually any domain and any game design, in order to understand how learning occurs and thereby help compare the results between studies. This would enable quantitative analysis of game properties for comparison. It would also provide a basis for objective measurement of learning outcomes by modelling the domain and comparing that to the learner’s mental model. Therefore, these modelling tools would allow us to study how to make better, more effective serious games.

Little is known about how best to design games for learning, or if certain domains of knowledge or skill necessarily lend themselves to a greater or lesser degree to game-based learning. Overall results have been somewhat mixed, with some successes and some failures. The central problem has been how to integrate the learning aims with the game design, because when they are separate what you end up with is “chocolate on broccoli” [9], [10]. Bogost [11] and Krajewski [12] explain that play should not be a reward for enduring the learning of unrelated knowledge.
and skills, but that the knowledge and skills should be empowering tools, instrumental to success in playing the game. That is the enigmatic potential of educational games – by having fun, people get the learning for free. Therefore, the educational potential of games should be found in the conditions where the learning is integral to the play.

We must take care to not present games as a panacea for all educational problems – surely they will have some apt areas of application, and some for which they are ill-suited. Why games? Why not a simple lecture or demonstration? The research on learning through exploratory play has revealed that the benefit of self-directed interaction is in learning about causation [13], [14], [15]. This makes sense, as learning causation is directly relevant to learning how you might be able to control and manipulate the system you are playing with – to attain mastery of the system. This is where the interactive aspect of games is most apparent and most potent. This is a way to avoid the “chocolate on broccoli” problem, by making the subject matter into the very game mechanics. If you instead just wanted students to be able to memorize a list of facts, then other learning techniques might be more effective. Once again, games cannot be a silver bullet for all education.

Therefore, the tool of formal modelling to use in game-based learning research should be one that is able to capture this kind of mastery of causal mechanisms through play. Games are fundamentally mathematical systems, and can therefore be analyzed using tools from complex systems theory. Virtually any real-world domain can also be modelled as a formal system, allowing direct comparison to (or transformation into) a corresponding game, and analysis using the same tools.

II. DYNAMIC CAUSAL NETS

In the aforementioned research on learning through play, the method of formal analysis is the causal Bayes net (CBN): An arrangement of nodes with links representing causation (fig. 1). These CBNs are used to represent the actual real world system, and also to represent the knowledge of a potential learner (the “belief net”) – which would start relatively empty and grow to (hopefully) match the CBN of the actual system. CBNs form an attractive starting point for game system modelling due to their usage in studies on learning through play, which involve learning causation through interaction with a system – exactly the kind of learning which is well-suited to games, thereby avoiding the chocolate-on-broccoli problem.

In a CBN, each node represents a possible state change (for example, a “weather” node could be rainy, cloudy, windy, or sunny), and each state can correspond to an arrow representing a causal influence (for example, the “grass” node could be dry or wet, depending partially on the weather). CBNs are very useful for trying to assess the probability of a certain state, given other states. For example, one might want to know, “what are the chances that the grass is wet, given the chance of rain today?”, or, “Given that the grass is wet, what are the chances that it rained today?” Causal Bayes nets are named so after Bayes Rule ($\Pr(B|A) = \Pr(A|B) \times \Pr(B)/\Pr(A)$), which tells us how to calculate the probability of an event given another event. There is no shortage of resources explaining the details of CBNs (see e.g. [14]), and so instead here I’ll discuss the advantages of slight modifications to this tool. CBNs are well-established and well-studied, but may require some modification for our purposes.

It is perhaps no coincidence that actual game design and analysis is frequently done using similar causal networks to those found in the cognitive science of learning through play (e.g. [16], [17], [18]). Causal networks seem to be an intuitive method for formally describing interactive systems to open them up to mathematical analysis. Many diagrams produced by game designers look very similar to CBNs that could have been made...
by cognitive scientists who might study learning in such games (figure 2). One major difference is that most cognitive science CBNs represent changeable states or events as nodes, whereas most game design networks represent “nouns” as nodes (a “noun” is simply any recognizable game entity, and it is contrasted with a “verb”, which is an interaction between nouns).

This shift of focus from states and events to nouns and verbs creates other differences in the modelling methods of traditional CBNs versus game design causal nets. For example, CBNs are always acyclic, meaning that the links of causation are not permitted to loop back around such that a state is able to (directly or indirectly) affect itself. In CBNs, the flow of cause and effect must always march forward towards the ultimate final state we are interested in, and never create feedback loops. This restriction is unacceptable for game systems, which almost always include cycles and feedback loops of some kind [16], as do real-world phenomena. By permitting cycles, the noun-based causal nets of game design become capable of simulating dynamic and chaotic systems. Therefore, to mark this distinction we will refer to these noun-based nets as Dynamic Causal Nets (DCN).

Comparing the subjects of analysis in these different domains reveals another critical reason for the difference. In most of the cognitive science research, the systems are kept relatively small and simple to control variables and so that they can be easily learned in the short timeframe of a laboratory experiment. When dealing with such small systems, it is practical and useful to represent state changes as nodes and probabilities as the links between them. In contrast, many of the game design diagrams are of large, complex game systems, involving subsystems that could be broken down even further. When dealing with such large systems, it is more practical to have the nodes refer to nouns, and let the state changes be implied by activation across the network.

Therefore, the dynamic causal networks of game design are able to keep their overall size and complexity more manageable by employing these conventions. The nodes represent different kinds of nouns, not specific instances of nouns. For example, there might be a noun to represent a dog, but not a noun to represent a particular dog, Rover, because Rover (being a dog) would have all the same links as every other dog. The links between nodes are the verbs, which describe the interactions between nouns. For example, maybe a dog has a 30% chance of eating a rabbit if it encounters one. This interaction would be represented with a link between those two nouns with a certain weight to convey the strength of that interaction (30%) and its negative effect (destroying the rabbit). In contrast, when a rabbit encounters another rabbit, it has a high chance of making a new rabbit. Therefore, the rabbit node would have a link connecting back to itself, with high strength and positive effect.

Further, the size of a node can be used to represent the density of that noun’s population in the world (e.g. a high population of rabbits in a small space), and thereby represent its likelihood to interact with other nouns. This variable, noun density, is the product of the number of the noun multiplied by the base rate probability. This base rate probability can represent physical size. For example, if the noun is “elephant”, it might only take 1 or 2 of them to have a high chance of encountering another creature in the same space, whereas if it is a small noun like a fruit fly, it might take 50 or more to achieve the same high probability of bumping into other creatures.

This kind of causal network (the DCN) is directly comparable to stock and flow diagrams from dynamic systems theory (e.g. [19], [20]), and to the system diagrams used in complex problem solving research (e.g. [21], [22]) – A branch of cognitive science also concerned with how people learn about complex dynamic systems, but not concerned with play. This allows for easier extrapolating of insights from research in the different fields of game-based learning, complex problem solving, system dynamics, etc. Achieving this parity would be difficult if one insisted on using traditional acyclic CBNs.

Of course, the links could represent any kind of verb, not just increasing and decreasing the populations or amounts of nouns. For example, they could represent spatial interactions, such as attraction and repulsion. Thereby, networks of nouns and verbs can be extremely powerful modelling tools.

Some of the most powerful systems of knowledge representation are simple at their core (See e.g. [23], [24]). They are composed of a small number of basic, abstract elements that can be combined in nearly infinite ways to model different subject matter. This power and broad applicability are useful properties of a modelling method, and an added benefit is that they can be easily turned into something that a computer can understand. For example, think of the binary system of a Turing machine. Therefore, an important factor to consider here is the low-level, abstract simplicity of a modelling method.

On the other hand, the goal of this modelling method is to be useful in educational game research. It should be a practical tool to help with the assessment of learning, and useful to game designers who need to model their domain and/or game. After all, the whole point of this exercise is to model game-based learning. For these purposes, there is value in having a recognizable high-level to the model so that the components are concrete objects, rather than abstract parameters or binary digits. For example, it would be considerably more complex and tedious for a human to model a domain or game in binary. Consider also when assessing learning, asking learners to join the nodes with lines of causation, they will better understand this instruction when the nodes are recognizable nouns, rather than abstract concepts. In that case, there are clear benefits to having
a modelling method that is able to be used to map at a high level that is intuitively understandable to a human. But we must be careful not to let all the implicit knowledge and inferences of human thinking to damage the formal precision gained from the modelling process.

This tension between the theoretical elegance and power of low-level abstractions, and the concrete, practical utility of high-level objects, is mirrored in the different fields that this modelling method hopes to bridge: The low-level mathematical domain of complex systems theory, and the high-level domains of cognitive science and game design. This is no small task, and the work is ongoing as we test this method’s robustness and versatility to model subjects in different domains. So far occasional adjustments have been found necessary to expand the versatility of this method, but the fundamental components have remained the same. We believe that the foundation of a dynamic causal net based on nouns as nodes and verbs as links is sound and will be powerful enough to serve the function needed in educational game research.

III. EXAMPLES

DCNs can be used to model virtually any real-world system or game, whether it is composed of multiple objects in discrete states, or a singular object with various continuous variables (such as an airplane). Fig. 4 is a simple example of a DCN.

Our work is ongoing in building a software tool for creating, simulating, and extracting data about these causal nets is currently in development. Once a DCN is made, it can be simulated in aggregate, allowing you to see the numbers for each node change in real time (for example, watching the population of rabbits go up). Or it can be simulated in detail spatially, where each node represents a population of agents moving and interacting in a 2D environment. For a model like an ecosystem, spatial simulation would be quite intuitive, whereas for a model of an airplane, it would make most sense to simulate in aggregate.

As noted in the introduction, game-based learning is not a panacea for all educational problems, and the proposed modelling method of DCNs reflects that, specializing in its ability to model knowledge, games or real-world domains that are amenable to game-based learning – learning causation through interaction. While it might not be able to represent every subject people hope to teach, it can represent a vast number of diverse subjects that are well-suited for game-based learning.

IV. SUITABILITY

DCNs have features that make them particularly well-suited to the study of game-based learning:

- Represent visible, interactive elements of systems, making DCNs ideal for analyzing the learning of causation and control via play with a system. Causal learning is the kind of learning most integrated with the play of games, as discussed in the introduction.
- Can be used to model virtually any real-world system or game.
- Capable of simulating dynamic and chaotic systems.
- Comparable to the nets used in various research fields, allowing easier interpretation and application of findings from these fields. This allows us to draw on broad and deep background research to derive predictions and hypotheses about game-based learning.
  - CBNs in the cognitive science of play
  - Stock and flow models in dynamic systems theory
  - Problem diagrams in complex problem solving
  - Design and analysis diagrams in game design
- Amenable to mathematical analysis and objective, quantified measures of complexity. This allows us to compare very different games according to their formal properties.

V. MEASURES & COMPLEXITY

A perfectly compatible measurement methodology to bring across from another domain, is how knowledge is measured in complex problem solving research. After the learning intervention, the learner is asked to connect the nodes of the system according to what they learned – which nodes affect which other nodes, in what way, and how strongly. This network they generate is compared to the actual system network and any links in common count as a “hit”, but any links different or missing count as a “miss”, creating a score of how well they learned the structure of the system. This method of assessing knowledge learned can be directly applied to DCNs, allowing researchers to quantify the amount of learning they achieved. These mental models, and the models of games and domains, can then also be analyzed in terms of their formal properties and complexity.
Since DCNs are dynamic systems, they may have chaotic properties such as sensitivity to initial conditions, and strange attractors [25], [19]. Sensitivity to initial conditions is a core indicator of complexity [26]. Network theory is rich with tools (see e.g. [27]) that can be used for quantifying the properties of causal nets. For example, measures such as the degree exponent, clustering coefficient, and diameter can provide general information on the structure of the causal network. Others such as overall connectivity, Platt index and the total walk count give a measure of the complexity of the network. Total walk count, for instance, is calculated by taking two nodes and counting the number of unique possible paths between them, and then going on and doing that for every pair of nodes in the network. The advantage of these network-based measures of complexity is that they are easy to understand in relation to the network-based measures of learning mentioned above.

But they are not the most universal measures of complexity. A very common general measure of complexity is algorithmic, or Kolmogorov complexity. This refers to the shortest possible description of something. For example, you may tell a story to a group of friends, and each one retells it differently while still conveying all the same information. There is theoretically a way to tell the story that minimizes the length of the story, while still containing all the same information. Kolmogorov complexity (KC) postulates that a more complex object would require a longer description than a simpler object. There are many valid and useful ways to conceptualize and measure complexity, and no doubt researchers will use those that best suit their individual context, but outlining them all is beyond the scope of this paper. Therefore, we will discuss how KC can be applied to DCNs.

The advantages of KC include its ubiquity and extensive study as a measure of complexity. Thus, its theoretical foundations and relationship to information and Shannon entropy are well established. It is also a universal complexity measure, not bound to a particular domain such as biology or physics. Anything that can be represented as a binary string can be analyzed in terms of KC, which means that causal nets that involve a different number of nodes, links, or additional components and parameters can be compared to each other, and to entirely different models such as actual simulation software, in terms of KC. This gives it a certain appeal as a universal currency of complexity.

KC was proposed by a Russian mathematician Andrey Kolmogorov in 1969 [28]. Even though the concept is named after Kolmogorov, several other mathematicians reached similar results about the same time independently [29]. KC provides a way to measure the description complexity of any objects. The description complexity means how complex it is to describe a given object. KC is approximate to entropy, the quantity of information [30].

For any object \( x \), let \( b \) be in the form of a string. Without losing generality, we suppose that every string in this section is a binary string with only “0” and “1”. Then let \( U \) be a Universal Turing Machine (A Universal Turing Machine can be informally considered as any modern computers with unlimited memory). Then the KC of \( x \) regarding \( U \) is defined as:

\[
K_U(x) = \min_{p:U(p)=x} l(p)
\]

where \( p \) is a program, also in binary string form, running on \( U \), which will print out \( x \) and then halt; \( l(p) \) denotes the length of \( p \). For all the programs with the same output \( x \), the KC of \( x \) regarding \( U \) is defined as the length of the shortest program, which will print out \( x \) and then halt when it runs on \( U \).

Even though the definition of KC depends on a particular Universal Turing Machine, it is universal because of the following theorem:

**Theorem 1** Let \( U \) be a Universal Turing Machine, \( \forall V \) which is another Universal Turing Machine, \( \exists c \) a constant, so for \( \forall x \in \{0,1\}^* \) (i.e. for each binary string \( x \))

\[
K_U(x) \leq K_V(x) + c
\]

The proof can be found in [30].

Due to the universality of KC, when we discuss the KC of an object \( x \), we can remove the reference machine \( U \); just denote the KC of \( x \) as \( K(x) \).

In the last few decades, KC as an objective complexity measurement has been applied in many different disciplines [31] including measure the complexity of very complex and abstract systems such as the design activities for a Global Software Development (GSD) project [32].

Even though KC provides a way to construct an objective measurement for many different types of complex system, theoretically, KC is not computable. In a real situation, no matter how long we spend on the problem, we can’t know for sure that there isn’t an even shorter, more efficient program that could do the job just as well. Therefore, we can only estimate an upper bound rather than give a precise measurement of KC. In this paper, we propose a scheme to estimate an upper bound of KC for a DCN.

Let \( N \) be a DCN, the proposed KC measurement contains two parts: the structure section and the parameter section:

\[
K(N) \leq K(N_s) + K(N_p)
\]

where \( N_s \) represents the structure information of \( N \) and \( N_p \) represents the parameter information of \( N \).

Suppose \( N \) contains \( n \) nodes with \( m \) directed connections. As there are possible total \( n^2 \) directional connections (we treat the connection from node \( a \), to node \( b \) and the connection from node \( b \) to node \( a \) as two different connections. We also consider a connection from a node to itself). Then the KC of \( N_s \) is estimated as:

\[
K(N_s) \leq K(N) + K(m) + K(\frac{n^2!}{m!(n^2-m)!}) \leq \log^* n + \log^* m + \log^* \frac{n^2!}{m!(n^2-m)!} + c
\]

where \( \log^* n = \log_2 n \); \( \log^* m = \log n + \log \log n + ... \) until the last positive term. \( c \) is a constant which can be ignored when we try to compare the KC of two objects. The proof of the above
formula is tedious and the detailed proof of a similar object (an \(n\times n\) matrix with all elements are 0 or 1) can be found elsewhere, in [33].

Suppose that \(N\) contains \(t\) independent parameters \(p_1, p_2, ..., p_t\) (please note we only need to count the independent parameters, some parameters can be directly calculated from other parameters; we will not include them in the KC estimation. For example, if we know the probability of \(raining\) is 0.2, we don’t need to record the probability of \(not\ raining\) as 0.8 because it can be directly calculated as \(0.8 = 1 - 0.2\).

Because each parameter is associated with one entity of network elements, which are \(n\) nodes and \(m\) connections, we need \(\log(n+m)\) bits to record the association information.

\[
K(N_p) \leq \log^* t + t \times \log(m + n) + \sum_{i=1}^{t} K(p_i) + c
\]  

(3)

where \(K(p_i)\) is the KC of the parameter. It can be estimated by the rules: if \(p_i\) is an integer, then \(K(p_i) \leq \log^* + c\) (here we ignore the trivial case of \(p_i = 0\)). If \(p_i\) has decimal places \(k\), then \(K(p_i) \leq \log^* k + \log^* (|p_i| \times 10^k) + 1 + c\). For example, if \(p_i = 0.875\), then \(k = 3\) and \(K(p_i) \leq \log^* 3 + \log^* (875) + 1 + c\).

Combining Eq(1,2, and 3), we have a way to estimate the upper boundary for DCNs. As a demonstration, we apply the formula to the DCNs shown in Fig 3 and 4 with the following results:

Example 1. DCN in Fig 3
\(n = 2, m = 2, t = 4, p_1 = 7, p_2 = 5, p_3 = 0.3, p_4 = 0.83\). Then

\[
K(N_p) \leq \log^* 2 + \log^* 2 + \log \frac{2^2}{2!(2^2 - 2)!} + c \leq 5 + c
\]

\[
K(N_p) \leq \log^* 4 + \log^* 4 + \log^* 7 + \log^* 5 + \log^* 3
\]

\[
+ \log^* 83 + \log^* 2) + 1 + 1 + c \leq 36 + c
\]

\[
K(N) \leq K(N_1) + K(N_2) \leq 41 + c\) (bit)
\]

Example 2. DCN in Fig 4
\(n = 10, m = 12, t = 9, p_1 = 0.4, ..., p_9 = 0.2\). Then

\[
K(N_p) \leq \log^* 10 + \log^* 12 + \log \frac{12^2}{12!(12^2 - 12)!} + c \leq 63 + c
\]

\[
K(N_p) \leq 124 + c
\]

\[
K(N) \leq K(N_1) \times K(N_2) \leq 187 + c\) (bits)
\]

From the above examples, we find that the KC of DCN in Fig 4 is about 4.5 times of the KC of DCN in Fig 3.

Through such measures, we would be able to compare different studies using different games and to test hypotheses. For example, it could be used to assess if there is an optimal level of complexity for a serious game to ensure the transfer of knowledge to the real-world domain it represents.

VI. DISCUSSION

It is often lamented that a standardized notation for formally describing game systems has yet to be established [34], [35].

Studies on serious games are most often tailored qualitatively to specific subject matter in an ad hoc way. This will need to change if research on game-based learning is to make significant headway. Without a unifying tool of analysis, it is difficult to design specific features of serious games according to research findings, it is difficult to draw broad generalizable conclusions beyond the subject matter of one study, and difficult to build directly on top of previous work if it has slight variations in methodology, subject matter or design.

DCNs offer a solution to this problem. Whether one’s interest is imparting knowledge, promoting behaviors, using games as motivational support during cancer treatment, or any other purpose, DCNs can be used as a common language of game structure to help discuss and compare game research. This study’s focus has been on how DCNs relate to one of the most central concerns of game design: Learning. DCNs can be used to model many different kinds of game design, real-world domain, or learner’s mental model. This allows quantitative measures of formal properties of different games and learning effects. The models are more easily comparable to those already used in other scientific fields, allowing research in serious games to draw upon work in the cognitive science of learning through play, system dynamics, complex problem solving, and game design. Most importantly, DCNs can be used to allow easier direct comparison of diverse serious game research, by giving objective and quantitative descriptions of differences in game design and subject matter. This allows experiments that explore these differences to make broadly generalizable conclusions about the design and knowledge transfer of serious games, allowing us to design better serious games in future.

Currently there is a critical shortage of knowledge on how learning improvements are linked to specific aspects of game mechanics. In sum, despite the extensive investment of time, money and infrastructure, that has been undertaken worldwide to roll-out games in classrooms, the knowledge base in this field both nationally and internationally is severely restricted. The power of games to instruct is recognized and theories that may help inform our understanding of how they influence learning are developing [36]. It is therefore very timely that knowledge on how to identify and manipulate the appropriate aspects of game mechanics and design are undertaken.

REFERENCES


