About the Intricacy of Tasks

Leonard M. Eberding,[1] Matteo Belenchia,[1,3] Arash Sheikhlar[1]
& Kristinn R. Thórisson[1,2]

[1] Center for Analysis and Design of Intelligent Agents,
Reykjavík U., Menntavegur 1, Reykjavík, Iceland
{leonard20, matteobel20, arash19, thorisson}@ru.is
[2] Icelandic Institute for Intelligent Machines
Reykjavík, Iceland
[3] University of Camerino
Via Andrea D'Accorso 16, Camerino, Italy

Abstract. Without a concrete measure of the "complicatedness" of tasks that artificial agents can reliably perform, assessing progress in AI is difficult. Only by providing evidence of progress towards more complicated tasks can developers aiming for general machine intelligence (GMI) ascertain their progress towards that goal. No such measure for this exists at present. In this work we propose a new measure of the *intricacy* of tasks, especially designed to describe their physical composition and makeup. Our *intricacy* is a multi-dimensional measurement that depends purely on objective physical properties of tasks and the environment in which they are to be performed. From this task intricacy measure, a relation to the knowledge of learners can allow calculation of the difficulty of a particular task for a particular learner. The method is intended for both narrow-AI and GMI-aspiring systems. Here we discuss some of the implications of our intricacy measure and suggest ways in which it may be used in AI research and system evaluation.

Keywords: Tasks · Environments · Intricacy · Difficulty · Task Theory · Artificial Intelligence · General Machine Intelligence · Evaluation · Training

1 Introduction

To better understand tasks and their role in research in general machine intelligence (GMI), we have been deepening our understanding of tasks and environments in the past few years, with an aim of developing a theory of tasks (cf. [15, 5]). This research has highlighted the requirement for proper analysis of tasks, including an objective measurement of a task's "complicatedness" or convolutedness. By this – and only by this – the difficulty of a particular task for a particular learning controller could be calculated, assuming that the difficulty of a task is a function of the task-environment (TE) and the controller performing the task. In this paper we introduce such an objective measure, called the ***intricacy*** of tasks, place it in the context of a causality-based task theory, and show the implications that such an intricacy measurement may have for progress in AI, and in particular, towards GMI.

Assuming there exist regularities in an agent's task-environment – which is a necessary requirement for any learning to be possible – these regularities can be expressed in the form of causal mechanisms. From these we may derive different measures of complexity[4] which can be used to calculate the level of intricacy of the task-environment. In past work we have described different complexity dimensions of tasks [14] and introduced an evaluation platform where these dimensions can be tuned by the analyst [7]. Our new approach to task intricacy is based on – and compatible with – this prior work.

The paper is structured as follows: In the first section we will show related work which indicates why such a measure of intricacy is of importance for AI research. Then we continue with causal principles of tasks as used to determine the level of intricacy of tasks. In the third section we place the intricacy measure in the context of agents and learning, describing its impact on the difficulty of tasks. Lastly we discuss the implications of this intricacy measure for AI, and conclude by listing some future work to be done using intricacy as a guide for better evaluation and AI system design.

2 Related Work

There exist different milestones in the history of artificial intelligence (AI) which were thought to have a decisive role in the research towards human-like, general machine intelligence (GMI). As the past has shown, most of these milestones did not necessarily lead towards more general AI systems. Solutions to the problems were rather more efficient and effective narrow AI systems. This discrepancy between expectations and actual results points towards the conclusion that the choosing of tasks for milestones might be flawed. The problem of choosing appropriate tasks for progress evaluation has been described before [1, 5, 7, 9, 14]. Each newly suggested milestone towards GMI systems can be argued against due to, for example, restricted context (Lovelace Test 2.0 [12]), human-centered approaches (e.g. Turing Test [17]), or too domain specific knowledge necessary for it (e.g. General Game Playing; cf. [13]).

One of the major evaluation platforms used nowadays – the Arcade Learning Environment (ALE) [4] – has been shown to have issues regarding the evaluation of progress of AI. Martínez-Plumed and Hernández-Orallo [10], for example, showed that some of the games of ALE do not indicate any progress towards better AI systems. This claim is supported by the findings of the developers of ALE: a brute-force tree search algorithm outperformed state of the art reinforcement learning algorithms in some of the games [4]. We argue here that the main reason for these problems is an insufficient understanding of the tasks themselves to fully understand the implications of agents solving different tasks. This coincides with the argument that tasks must be analyzed more thoroughly

[4] With complexity we mean the intuitive concept as used in every-day language, not the concept as used in computer science.

to support progress in the field of AI [8, 9, 15]. The SAGE (Simulator for Autonomy and Generality Evaluation) platform was developed particularly for this purpose and has shown the advantages a deeper insight into the task's complexity dimensions can have for the evaluator [7]. But again, there currently does not exist a measure of difficulty of the tasks presented to the agents and no measure of change in difficulty, if complexity dimensions are adjusted.

3 Causal Principles of Tasks

The aim of general machine intelligence (GMI) research is to create systems which are able to cope with highly complex worlds, like the physical world, and to be able to do a multitude of (unrelated) tasks in these highly complex environments. For this regularities of the world must be learned and knowledge about the environment accumulated. These regularities can be seen as 'mechanisms', representing functions which determine the value of effect-variables by using the values of cause variables. This knowledge of any learner is the result of a composition process, pieced together incrementally from experience with the world over time, accumulating in a semi-systematic way. This kind of learning is a constraint on any autonomous learner that doesn't have complete information at birth. If an agent is to learn independently, without help from teachers or some other source, its knowledge acquisition processes must be self-guided—it must have a capacity for *cumulative autonomous learning*. Achieving goals in the context of any phenomenon necessarily requires knowledge of causal relations, in particular of the causal relations that relate manipulatable and observable variables of the phenomenon to the goals of an assigned task. The existence of any causal relations between relevant variables must either be known by a performing agent or discovered by it in the process of performing a task.[5] These causal-relational models therefore are at the center of any task description. The intricacy measure introduced here relies specifically on these models using their interconnections as a measure of the "complicatedness".

Aside from the internal cause-effect-structures of the environment, the body of the agent, including sensors and actuators, must be taken into account. The noise that can take place when measuring/observing variables and interacting on them needs to be modeled in the causal structure of the tasks. Therefore, manipulatables and observables are treated differently to other internal variables of the environment. Variables which are theoretically measureable are the causal parents to the actually observed variable, which is used as input to the controller, and includes observation noise. Manipulatable variables of the environment, on the other hand, are causal children of the controller's chosen action (again, including the noise of actuators).

[5] For a more detailed description of our understanding of causal knowledge and its implications see [3].

From the causal connections between causes and effects – as in "A leads to B" – causal relational models can be derived.[6] In the following sections we adapt the notion of such causal relational models for the purpose of obtaining an objective measure of a task's "complicatedness" (complication) based on this kind of models. This is the most fundamental assumption that we can make about any task-environment in which learning is possible: The assumption of the existence of causal relations—*AECR:* Only then, prediction, planning, and directed interventions are possible, and only then tasks can be executed at all.

3.1 Causal-Relational Models

Causal-relational models (CRMs) are representational construct for general learning. CRMs encode actionable information, in the sense that they can be used to get things done (taking action with foreseeable results), predicting future states, derive the causes of observed events, explain observed phenomena, and act as re-creation of the causal relation [16]. The kind of models that we are talking about are *causal-relational bi-directional* models, where by bi-directional is meant that they can be used in forward-chaining to produce predictions of future states and in backward-chaining to produce paths towards goals. By causal-relational is meant that they encode procedural (causal) knowledge, where the left-hand side (LHS) is a pattern representing the cause and the right-hand side (RHS) is a pattern representing the effect. The CRMs represent a relation between the two patterns such that we can forward- and backwardchain from causes to effects and vice versa. Additionally there exists a separate set of the required conditions under which the relation between LHS and RHS holds, thereby specifying in which situations a certain CRM is relevant. The RHS represents the post-conditions of the LHS pattern. In forward-chaining, when the LHS pattern is observed, a prediction based on the RHS can be generated by a process of deduction. In backward-chaining, when the RHS pattern is observed and it is a goal, a sub-goal based on the LHS can be generated. Sub-goals can be further backward-chained until a manipulatable variable is reached. This way models can be used to produce effective plans to achieve goals and help to analyze tasks for their inner causal structures including manipulatables, observables, goals, and sub-goals. Causal relational models are therefore ideal to be used as the underlying principles of intricacy. They describe the task fully, and give insight into what needs to be known by an agent to perform well in the task (i.e. observe the environment, do correct planning, and take actions to achieve a goal).

3.2 Causal diagrams

A *task* can be described, from its designer's perspective,[7] as weakly-connected causal chains. When this is done, a task is reduced to a deterministic form

[6] While we take the non-axiomatic approach we still assume that the underlying environment follows certain rules, i.e. causal structures.

[7] We assume that the "designer's perspective" includes a complete access and overview to a task's full set of variables.

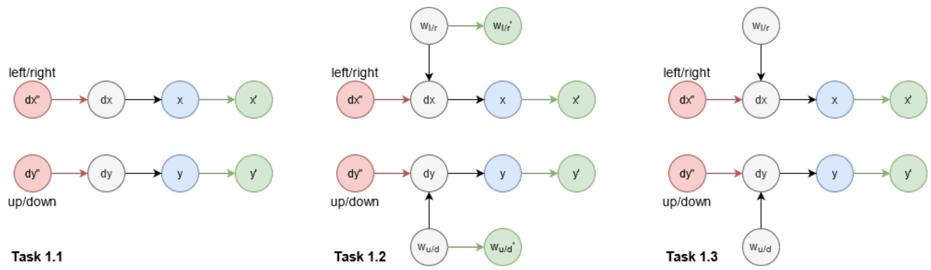

Fig. 1. Three examples of different levels of intricacy for similar tasks. Goals are to reach a certain X/Y position on a grid. The learning/performing agent can execute actions of moving left/right or up/down. Colors in all tasks: *Red:* actions as executed by the controller; *Green:* Observables as inputted to the controller (including observation noise, if applicable); *Blue:* Goal variables; *Grey:* Other variables. *Task 1.1:* Task of moving to a certain position in an open space; *Task 1.2:* Task of moving to a certain position with walls which can be seen; *Task 1.3:* Task of moving to a certain position with invisible walls. The level of intricacy rises from left to right. Colored arrows indicate transitions of data between controller and environment

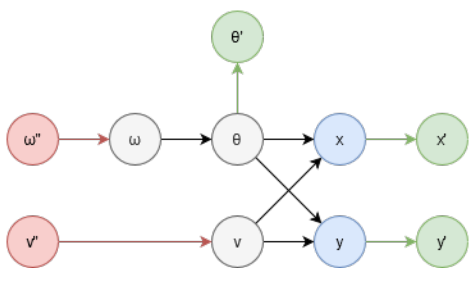

Fig. 2. Task 2: A more complex example of a task represented using causal-relational models. The goal is to reach a certain X/ Y position. The environment is continuous rather than a grid and the agent has the control to either turn on the spot or move forward/ backward. Same color coding as in figure 1.

that can be represented by the bi-directional models, capturing the whole task's dynamics. Additionally, inaccuracy of actions and measurements must be taken into account. Therefore, in this description of tasks, variables are not directly observable or manipulatable but instead a noisy causal child or parent acts as an observable or manipulatable, respectively. In figure 1 three similar tasks are shown. The task is to move to a certain goal position inside a grid-world. Figure 2 shows the same goal in a continuous world where a learning/performing agent can rotate on the spot or drive forward. Intuitively speaking, it is easy to describe the "complicatedness" of these four tasks. With larger causal relational model-networks this becomes a much harder problem. It is for this reason that we propose our measurement of *intricacy*.

4 Intricacy

We define the *intricacy* of a task as the measure of a task's "complexity" based purely on objective parameters. This way we can rank tasks by their intricacy and have an objective way to assess the progress of AI systems. Additionally, our notion of intricacy gives other implications for the design of GMI-aspiring systems. The definition and implications are presented in the following section.

4.1 Definition

Intricacy is an objective multi-dimensional measure consisting of the following physically-based, measurable properties of a task (ordered by their weight on the intricacy value):

1. **The minimal number of causal-relational models needed to represent the relations of the causal structure related to the goal(s).** This minimal number of models is an objective measure which depends solely on the particular specification of the task (inclusive of the controller's body), and captures all the relevant parts of the task proper, leaving out possibly unnecessary details and relations that are superfluous for the task. This means that, for example, tasks which contain superfluous variables and relations have the same intricacy of the same task abridged of all superfluities. The steps to obtain this minimal number of models entail the identification of all the relevant causal chains, turning them into relational models and then quantify them. This additionally means that a task's intricacy is dependent on the level of detail in which the task is performed.[8]
2. **The number, length and type of mechanisms of causal chains that affect observable variables on a causal path to at least one goal.** The three parameters of concern to this definition are (a) how many distinct causal chains there are, (b) how long are they in terms of number of variables involved and, (c) the complexity of the functions that define the mechanisms on these causal chains. As a measure for this dimension we suggest the Vapnik-Chervonenkis dimensions [6] or the Rademacher complexity [2], which includes the Vapnik-Chervonenkis dimension bound. (In the context of statistical learning, the class of functions with a lower Rademacher Complexity can be understood to be easier to learn.)
3. **The number of hidden confounders influencing causal structures related to the goal.** Other things being equal, hidden confounders should make the learning of previously described relational models of causal structures much harder. Therefore we include the number of unobservable variables influencing goal-related causal structures and chains in our intricacy measurement.

Intuitively, the intricacy of a task is a measure of what physical mechanisms are in place that need to be known by any intelligent being, whatever is its architecture, knowledge or capabilities, to perform the task in the given environment (inclusive of the controller's body). The task's intricacy is invariant on the initial values

[8] For further information on the level of detail see [3]. How knowledge representation of the agent affects the intricacy by changing the level of detail is a problem that still needs to be addressed.

Task	Relational Mdl.	Causal Mech.	Confounders
1.1	2	2, linear	0
1.2	4	4, linear & non-linear	0
1.3	4	4, linear & non-linear	2
2	5	5, linear & non-linear	0

Table 1. The four different properties describing the intricacy of tasks for the given task examples from figures 1 and 2. As expected the results show, that the task in continuous space (Task 2) is the most, and the grid-task in an open space (Task 1.1) is the least intricate.

of the task's variables. From the definition of intricacy it follows that the higher the intricacy, the lower the size of the solution space[9] and vice versa.

Coming back to figures 1 and 2 we can now argue the different levels of intricacy of the four tasks (see table 1). For simple tasks as shown in the two figures the level of intricacy is easy to determine intuitively. For more complex tasks such a measure, however, becomes more important due to its implications for the evaluation of AI systems.

To use this measure of intricacy for AI evaluation, or as a support in AI system design, it needs to be connected to the learning agent. For this we introduce the *effective intricacy*, which not only takes task features into account but also connects the task to the experience of the learner.

4.2 Effective Intricacy

The effective intricacy of a task is an agent-dependent version of intricacy, as defined above, where an agent's previously acquired knowledge that it brings to the task is taken into account. Effective intricacy is thus a measure of intricacy minus any intricacy that is known by the agent, and thus made irrelevant to the computation of difficulty. It uses almost the same properties of the task-environment as the intricacy measure. However, the effective intricacy only depends on **unknown** (to the agent) properties.

1. The minimal number of **unknown** causal-relational models needed to represent the causal relations related to the goal(s).
2. The number, length and type of **unknown** mechanisms of causal chains that affect observable variables on a causal path to at least one goal.
3. The number of **unknown**, neither directly, nor indirectly (through causal children) observable variables directly influencing the causal chain(s) between manipulatables and goal variables.

4.3 Difficulty

While effective intricacy represents the physical aspects of a task that are relevant to how difficult it may be for a particular agent, the difficulty of a task

[9] For a more detailed view on the solution space of tasks see [15, 3].

includes the agent's ability to learn relational models. Additionally, the precision of the agent's transducers and available resources – including time and energy – must be taken into account (especially if assuming the assumption of insufficient knowledge and resources (AIKR; [18])). Additive noise, for instance on observations and actions, can make a task more difficult. When actuators do not generate reliable interventions, or sensors reliable observations, the usage of causal models becomes more unreliable, making accidental mistakes possible. The difficulty D can therefore be expressed as the cross product of controller C and task-environment TE: $D = (C \times TE)$. Or more precisely the task's intricacy I, resources R, and transducer noise N_T: $D = (C \times I \times R \times N_T)$.

5 Learning & Performing

Aside from the possibilities an intricacy measure opens for AI evaluation, it also brings strong implications for other areas of AI research including the learning and doing of tasks, and the design of AI systems.

Learning a task. The process of learning a task can be thought of as the search for relational models that can bring about a *satisficing* solution to the task. This search for models is driven by looking for associations in observable variables, finding which of these associations is of causal nature and growing the understanding of how each of these variables map onto the goal variables.

- By learning the causal structure of a task, a learner decreases the **effective intricacy**, since that knowledge allows it to find effective ways of controlling and achieving the goals. The more spurious associations are removed, the more useful the causal-relational models and thus, the lower the effective intricacy becomes.
- The importance of variables is revealed when the causal relations are discovered. Reciprocally, detecting the important variables enables the learner to find causal relations that are useful for performing a task and therefore reduces the effective intricacy.
- If a learner discovers all causal relations in the state space (without taking the importance of variables into account), changing the goals does not affect the effective intricacy, since the learner is already aware of how to conduct a new task within the same environment.
- When the learner is aware of the complete causal structure of a task, the deadline of the task and the energy required to perform it become the decisive measures for difficulty.
- When the learner knows all causal structures of the task-environment and has sufficient available resources, the only remaining part of the computed difficulty is the noise in the transducers.

Performing a task. A good controller of a system (performer of a task) is one that already knows how to achieve the task under a range of environmental conditions. The agent's performance on the task allows us to draw conclusions on the effective intricacy, and the difficulty of the task.

- To a controller that performs a task perfectly repeatedly, the effective intricacy is zero.
- The higher the effective intricacy of a task – or the lower the amount of experience related to the task – the more difficult a task becomes to do, other things being equal.
- If the effective intricacy is equal to the task's intricacy an agent must rely on random interactions until it has learned enough, reducing the effective intricacy.

6 Conclusion

In this paper we introduced a new measurement to describe the "complicatedness" of tasks [3]. With our intricacy measure we are able to describe the effective intricacy, producing a concrete definition of task difficulty in relation to GMI-aspiring agents. We believe that through this measure, a more sophisticated choice for tasks to evaluate and compare AI systems is possible.

While we have provided evidence for the usage of this intricacy measure there still is work to be done to automatically calculate the intricacy of a wide variety of different tasks to evaluate the scalability and applicability of our approach. For this, the GMI aspiring system AERA (Autocatalytic Endogenous Reflective Architecture) [11] could be adapted, as it already provides the ability to extract causal relational models by interacting with the world. It could therefore be a good starting point for automatic intricacy calculation of more complex tasks. Another future idea would be to calculate the intricacy values of different tasks of the Arcade Learning Environment (ALE) and compare the results with the conclusions drawn by [10] using Item Response Theory (IRT) to determine the usefulness of different ALE tasks for progress evaluation.

From there we hope to be able to draw a connection between the intricacy of tasks, which an agent is able to solve, and the system's generality. This would provide researchers an additional measure of generality independent of the task-environments used for evaluation.

Acknowledgments. This work was supported in part by Cisco Systems, the Icelandic Institute for Intelligent Machines and Reykjavik University.

References

1. Adams, S., Arel, I., Bach, J., Coop, R., Furlan, R., Goertzel, B., Hall, J.S., Samsonovich, A., Scheutz, M., Schlesinger, M., et al.: Mapping the landscape of human-level artificial general intelligence. AI magazine **33**(1), 25–42 (2012)
2. Bartlett, P.L., Mendelson, S.: Rademacher and gaussian complexities: Risk bounds and structural results. Journal of Machine Learning Research **3**(Nov), 463–482 (2002)

3. Belenchia, M., Thórisson, K.R., Eberding, L.M., Sheikhlar, A.: Elements of task theory. In: Proceedings of the International Conference on Artificial General Intelligence. Springer (2021, in submission)
4. Bellemare, M.G., Naddaf, Y., Veness, J., Bowling, M.: The Arcade Learning Environment: An evaluation platform for general agents (extended abstract). In: Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI). p. 4148–4152 (2015)
5. Bieger, J., Thórisson, K.R., Steunebrink, B.R., Thorarensen, T., Sigurdardóttir, J.S.: Evaluation of general-purpose artificial intelligence: Why, what & how. EGPAI 2016 - Evaluating General-Purpose A.I., Workshop held in conjuction with the European Conference on Artificial Intelligence (2016)
6. Blumer, A., Ehrenfeucht, A., Haussler, D., Warmuth, M.K.: Learnability and the vapnik-chervonenkis dimension. Journal of the ACM (JACM) **36**(4), 929–965 (1989)
7. Eberding, L.M., Sheikhlar, A., Thórisson, K.R., Andrason, S.P.: Sage: Task-environment platform for evaluating a broad range of AI learners. In: International Conference on Artificial General Intelligence. Springer (2020)
8. Hernández-Orallo, J.: The measure of all minds: evaluating natural and artificial intelligence. Cambridge University Press (2017)
9. Hernández-Orallo, J., Baroni, M., Bieger, J., Chmait, N., Dowe, D.L., Hofmann, K., Martínez-Plumed, F., Strannegård, C., Thórisson, K.R.: A new ai evaluation cosmos: Ready to play the game? AI Magazine **38**(3), 66–69 (2017)
10. Martınez-Plumed, F., Hernández-Orallo, J.: Ai results for the atari 2600 games: difficulty and discrimination using irt. EGPAI, Evaluating General-Purpose Artificial Intelligence **33** (2016)
11. Nivel, E., Thórisson, K.R., Steunebrink, B., Dindo, H., Pezzulo, G., Rodríguez, M., Hernández, C., Ognibene, D., Schmidhuber, J., Sanz, R., Helgason, H.P., Chella, A., Jonsson, G.K.: Bounded recursive self-improvement (2013)
12. Riedl, M.O.: The lovelace 2.0 test of artificial creativity and intelligence. arXiv preprint arXiv:1410.6142 (2014)
13. Świechowski, M., Park, H., Mańdziuk, J., Kim, K.J.: Recent advances in general game playing. The Scientific World Journal **2015** (2015)
14. Thórisson, K.R., Bieger, J., Schiffel, S., Garrett, D.: Towards Flexible Task Environments for Comprehensive Evaluation of Artificial Intelligent Systems & Automatic Learners. In: Proceedings of the International Conference on Artificial General Intelligence. pp. 187–196 (2015)
15. Thórisson, K.R., Bieger, J., Thorarensen, T., Sigurðardóttir, J.S., Steunebrink, B.R.: Why artificial intelligence needs a task theory. In: International Conference on Artificial General Intelligence. pp. 118–128. Springer (2016)
16. Thórisson, K.R., Talbot, A.: Cumulative learning with causal-relational models. In: International Conference on Artificial General Intelligence. pp. 227–237. Springer (2018)
17. Turing, A.M.: I.—Computing machinery and intelligence. Mind **LIX**(236), 433–460 (10 1950). https://doi.org/10.1093/mind/LIX.236.433, https://doi.org/10.1093/mind/LIX.236.433
18. Wang, P.: Rigid flexibility: The logic of intelligence. Springer Science, vol. 34 (2006)