A Roadmap for Causality Research in Complex Adaptive Systems

Samuel Reyd, Ada Diaconescu, Jean-Louis Dessalles
Telecom Paris, LTCI, IPParis, Palaiseau, France
name.surname@telecom-paris.fr

Lukas Esterle
Aarhus University & DIGIT, Aarhus, Denmark
name.surname@ece.au.dk

***Abstract*—Causal modeling and cause-detection (abduction) are critical challenges for understanding and managing complex adaptive systems (CAS), including self-improving and self-integrating systems (SISSY). While the vast research literature on causality addresses many important issues, it mostly lacks specific methods for dealing with CAS-related issues – e.g. runtime changes and uncertainty; integrating new types of variables on-the-fly; non-linear phenomena and large system scales. This paper aims to identify some of these key challenges and to highlight relevant research fields that might help to address them.**

I. Introduction

Causal models and associated abduction methods are essential for system understanding and management based on rational, predictive and experience-driven processes. Yet, most common causal modeling and analysis techniques (e.g. Bayesian networks [1] and corresponding abduction and negation methods [2]) are highly generic, assuming key modeling aspects as given, trivial or complementary – e.g. modeling scope, abstraction level(s), known system variables and relevant time intervals, or statistically-relevant observability in repeating conditions. These issues represent blind spots for most causality research, which focuses instead on identifying 'flat' causal models (e.g. directed acyclic graphs (DAGs) at a given scale), based on predefined variables, within an implicitly defined scope and timing scheme [3]. While highly relevant to closed systems running in relatively stable environments – e.g. 'classic' engineering applications, 'traditional' risk-assessment economic models, or static neural networks – such contributions provide little insights for handling critical issues of complex adaptive systems (CAS) or self-integrating and self-improving systems (SISSY) – e.g. scalability; open adaptability to unpredictable events; complex, chaotic or emergent causality relations. This paper aims to highlight such blind-spots in the 'mainstream' causality literature and provide a road-map for future efforts that would alleviate this situation. It encourages focus on key causality concerns and motivates knowledge transfer from relevant areas – e.g. complex, adaptive, multi-scale system modeling. While proposing promising directions for future research, the paper does not aim to specify a concrete project, implementation method or support platform; nor does it claim to guarantee a particular outcome.

In brief, section II summarises 'mainstream' causality-related approaches. Section III highlights key outstanding challenges when applying these approaches to CAS. Finally, section IV proposes a road-map for addressing these challenges by integrating solutions from CAS-related domains. Figure 1 summarises the main issues and research directions.

II. Classic approaches to causality

Understanding causal relations in "classic" engineered systems may help with system optimization, risk assessment, prediction and overall problem-solving. These abilities are also crucial for CAS and specifically SISSY. Such systems can utilize causal relations for (self-)explanation, prediction and adaptation, i.e. knowing why something happens to get insights into the system's functioning and adapt its behaviour.

Obtaining a cause first requires a causal model that expresses causal links between variables. Then, an abduction method is needed to search for causes within the model. The cause must describe the influence of the cause variables' values over the effect. While modern causality research is rich and broad, it can be categorised into two main directions: (i) building causal models; and (ii) performing abduction within a causal model. Such research typically targets systems deployed in static environments where all relevant variables are known in advance and feature relatively tractable causal relations. In this paper we highlight several additional challenges specific to CAS that evolve in dynamic, unpredictable environments.

A. Causal discovery

The field of causal discovery focuses on finding causal links between a set of given variables. The goal is to find a Causal Bayesian Network (CBN) that captures relations between these variables. A CBN is a Direct Acyclic Graph (DAG) where the nodes are the variables and directed edges are causal links between variables. For instance, $X \rightarrow Z \leftarrow Y$ is a CBN for variables X, Y, Z where X and Y (the *cause variables*) have a causal influence on Z (the *effect variable*).

Causal discovery aims to classify statistical links between variables in a dataset – e.g. direct causal links, cofounders, uncertain causality. The focus is neither on how these causal links came about nor on how a cause's values may influence an effect's behaviour. As obtaining a DAG from a dataset is not always possible, other types of graphs are defined to represent available information – e.g. Completed Partially Directed Acyclic Graphs, Acyclic Directed Mixed Graphs, Maximal Ancestral Graphs, Partial Ancestral Graphs, and many others. Obtaining these graphs requires various types of assumptions,

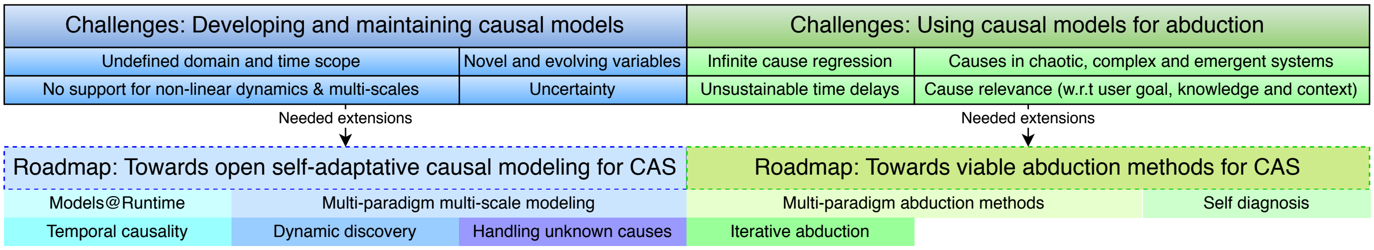

Fig. 1. Summary of identified challenges and promising research directions for causal modeling and abduction methods for CAS

such as d-separation, d-faithfulness, sufficiency, m-separation, etc. CBN-based approaches increased in popularity with the development of machine learning [4] and big data [5].

B. Actual causation using structural causal models

Pearl [1] offers an intuitive and mathematical description to answer 'Why did X happen?'. A cause is defined as an event (e.g. a variable taking a value) that, if not true, would have led to an outcome different from the one observed (e.g. a change in the effect variable's value). Such cause-effect relation between variables is expressed via a Structural Causal Model (SCM). A SCM consists of a set of variables and functions, called structural functions, which set the values of each variable given the values of the others, e.g. $Z = f(X, Y)$ in the example above. Notably, SCMs support 'interventions'– i.e. forcing a variable's value regardless of the others. This operation allows counterfactual thinking and avoiding spurious correlations. Extensive research has further developed this definition [6].

The SCM formalism deals with how causal influence operates. It is general enough, theoretically, to express any cause, even within CAS or SISSY. However, the approach as such does not include ways to discover, to represent or to modify a SCM. While some recent approach may aim to achieve this, we are unaware of any systematic mainstream proposal to adapt SCMs to deal with open, complex, self-* systems.

III. Causality Challenges for CAS

As CAS are rarely considered in current causation literature, applying common approaches to CAS requires significant effort. We hereby highlight several challenges one must face to develop and maintain SCMs for CAS (III-A); and suggest promising venues for addressing them. We then identify further issues with SCM usage for abduction methods (III-B).

A. Developing and Maintaining SCMs

Actual causation requires a SCM that consists of: i) a predefined set of variables; and, ii) structural equations interrelating these variables (i.e. computing variable values from other variable values). According to Pearl [1] causes in such SCMs are expressed as causal variables taking certain values, e.g. effect $Z = z$ happened because variable X took value x. Hence, causal variables should represent modeling concepts for the targeted system. Some variables should be suited for modeling system dynamics (e.g. cause-effect phenomena at a technical level); and others for providing explanations to users (e.g. conceptual level). In addition, structural equations specify how the values of inter-dependent variables impact each-other. SCMs represent these via generic functions (e.g. $Z = f(X, Y)$). This requires considerable expertise to express non-trivial dynamics, as needed in CAS modeling.

These two key aspects of SCM are hardly straightforward to determine in large, complex, adaptive systems due to challenges in two areas: (1) determining relevant variables; and (2) identifying and representing their underlying relations.

*1) **Relevant variables and their dynamics:*** In the context of CAS, or SISSY, identifying and adapting the suitable set of variables for causal modeling poses significant challenges.

Domain scope. The parts of the system and its environment that are relevant for causal modeling and analysis are often implicit and decided statically, once and for all. Even if causal discovery allows for latent variables [1], the selection of relevant variables is usually not included in the modeling process. Moreover, for large systems that adapt frequently within complex environments, maintaining the entire causal model may incur unbearable overloads. E.g. smart home models typically include the observable variables of all smart devices, plus a few external indicators (e.g. temperature, power prosumption, cost [3]). When an electric device malfunctions, the cause may be an internal component, a defective wall plug, a house blackout, or a broken power wire several kilometers away from the house. In this case, how far should the scope of the causal model stretch out to – include each device, the entire house, the surrounding neighbourhood or the entire world? The variables included in the model should include all plausible causes while meeting limited resource constraints.

Addressing this challenge requires developing causal models on-the-fly, extending and shrinking their modeling scopes as needed to answer particular questions. Specific abduction methods are needed to adjust modeling scopes to each question, in an iterative manner, depending on the plausible causes found within each iteration. Feedback from user dialogue and other verification processes (e.g. self-testing and -diagnosis) may be adopted here.

Conceptual abstraction level. Variables may be defined at different abstraction levels, or scales, based on their observation granularity, in space and time [7], [8]. E.g. direct observations (e.g. a thermometer's temperature) may be aggregated into higher-abstraction concepts (e.g. the average temperature of an entire house, or over an entire week; or the user concept of 'warm' [9]). These can be represented as specific (sets of) variables. However, certain abstractions may represent emergent phenomena that result from the interactions and accumulated effects) amongst individuals and hence cannot be assigned to single entities but only to an interrelated collective (e.g. cellular automata dynamics [10] or swarm movement [11]). As some emergent phenomena are unpredictable offline, causal models must discover novel abstract variables at runtime. Moreover, while an effect may be explained via several alternative causes, the relevant cause may pertain to a particular abstraction level – this depends on the stakeholder's purpose, knowledge and context. E.g. when a driver asks why their car moves slowly the relevant answer is that they are in a traffic jam, and not that the car in front moves slowly [12]. Hence, determining suitable abstraction level(s) for relevant cause(s) is a difficult, yet essential issue [11].

Feedback loops. Most causal models rely on directed acyclic graphs (DAG) [1], [3]. When causal traces form cycles (e.g. feedbacks), finding relevant causes becomes more complicated. E.g., in a smart home, the relevant cause for an observed temperature is usually not the temperature that preceded it. Instead, it may be the user input to a thermostat and the continuous feedback loop that this engenders between its thermometer and heating system, over a certain period.

As feedback loops are rather the rule than the exception in CAS, they should be modelled as first class entities, or variables, and employed as such to provide relevant causes. This issue is included within the wider concern of abstraction level, as an entire loop may be abstracted as one higher-level variable. Moreover, feedback loops may also occur between abstraction levels [13], [7], further complicating the identification of the appropriate abstraction level(s).

Novel and evolving variables. In open, self-* systems, the complete set of relevant variables cannot be known in advance or determined statically once and for all. CAS may need to handle adding a new device type to a smart home; discovering an emergent phenomena, e.g. feedback between an automatic window and a thermostat; or adjusting the user's concept of 'warm' to a temperature range that shifts between seasons.

Hence, causal models need to discover and integrate new types of relevant variables and to update their representations when they evolve over time. Verifying the correctness of such changes to the causal model raises additional challenges.

Unknown unknowns and limited information. As CAS commonly operate within complex, dynamic environments, it is impossible to know all relevant parameters, components, or other systems, at any one time. This is especially the case when systems are ill-equipped to observe certain variables, which seemed irrelevant in the system's past. Due to limited information, verifying cause and effect can become challenging. If an unknown system is causing something, this cannot be modelled explicitly based on known variables [14].

New modeling concepts must be created dynamically to make-up for the observed effects of unknown variables. Such concepts must be refined over time, as new observations become available (cf. 'evolving variables' issue above).

*2) **Variables relations and their dynamics:*** In the context of CAS, accurately representing the interrelations between variables raises significant issues. Relevant expertise has been developed for CAS modeling, including non-linear phenomena, e.g., phase transitions, emergent properties or chaotic behaviour. To extend current generic SCMs, we highlight classic notions of systems dynamics and suggest integrating them within SCM representations to support more effective abduction processes and produce more relevant causes.

Phase transitions and bifurcation points. Non-linear system dynamics is often regulated by feedback loops – with negative feedback having a dampening, stabilising effect; and positive feedback leading to super-linear, exponential growth. CAS regulation often fluctuates between negative and positive feedback. This leads to alternating periods of stable behaviour within a single attraction basin and exponential growth periods shifting behaviour from one attractor basin to another (i.e. punctuated equilibrium) – e.g. exponential economic growth punctuated by crises and stabilising plateaus; or exponential population growth punctuated by eco-system collapse.

Such alternating dynamics have been studied and represented via bifurcation theory (where small changes in a system's dynamics may lead to different trajectories, within different basins of attraction); or phase transitions (where a small change in a quantitative system variable may lead to a qualitative difference in a systemic property and/or behaviour) – e.g. the transition of ice to water to steam at varying temperatures, and the ensuing systemic properties. Such representations may prove essential to causal modeling, as events leading to bifurcations and phase transitions provide key causes for observed changes (effects) in system dynamics and global properties. E.g. high temperatures crossing specific thresholds provide relevant systemic causes for forest fires, even though specific events such as a spark may be the actual triggering event. Similarly, the relevant cause of an economic crisis may be a Ponzi scheme or speculative bubble, whereas the specific trigger may be linked to an individual action.

Time scope and temporal abstraction level. The time scope establishes the interval over which the system dynamics model should expand (e.g. for a predefined period or until a stopping condition is met). The abstraction level determines the granularity of time representation (e.g. continuous time, discrete events or time series at various frequencies). These time-related aspects are key to modeling system dynamics. The related difficulties are somewhat equivalent to those of the domain scope and variable abstractions highlighted above.

B. Using SCMs to Search for Actual Causes

Actual causation has generated a wide range of contributions, concerning e.g. causal sufficiency [15] and responsibility attribution [16]. Still, finding relevant causes for phenomena observed within complex environments raises significant problems. Hence, most formal abduction approaches necessarily focus on relatively simple, illustrative use cases. We highlight here several key concerns to be addressed when dealing with non-trivial dynamic systems, such as CAS.

The abduction process is typically triggered by an observer upon noticing a problem, a surprise or a conflict [2] – i.e. a discrepancy between an observed and an expected or desired situation. The observer questions the conflict to gain better understanding and/or to solve it.

*1) **Computability and Usability of SCM**:* While generating causal models already raises difficult issues (subsec. III-A), maintaining and employing them with limited computing and time resources raises additional challenges.

Infinite cause regression. Finding a cause to a problem may not satisfy the observer – e.g. they consider the cause irrelevant (cf. issue of 'relevant causes' below); or the cause raises a further question. In this case, another search may be launched to find an upstream cause. This process may repeat recursively, raising the question of when to stop.

Recursive abduction methods, such as Conflict-Abduction-Negation (CAN) [2] address this requirement, ensuring relevance via iterative dialogue with the user (observer). The process ends when the user asks no further questions. This, of course, does not guarantee satisfaction (e.g. the user may merely get frustrated about not getting a relevant answer after a while and give-up the process). Alternative approaches must be adopted in the absence of human observers (e.g. unmanned search-and-rescue or space systems). These may include limited causal-chain iterations and system self-diagnosis.

Causes in chaotic systems. A key characteristic of a chaotic system is that small changes in one system state can result in significantly large differences in later states (e.g. a double pendulum system, or the 'butterfly effect'). This implies that for finding the cause of a system state at time t it makes little sense to analyse its previous individual states at $t-1$, $t-2$, .. $t-i$ in isolation.

Causes in complex systems. Identifying relevant cause(s) for a problem in CAS may be challenging due to the system's high number of variables and their non-linear interrelations – e.g. economic bubbles, forest fires or climate change. Moreover, the relevant cause(s) may not be the most obvious ones, but rather the ones that trigger a chain of events leading to the observed problem. In other cases, the problem is due to the accumulated effect of an event sequence, where each event was necessary but insufficient for causing the problem (e.g. as revealed in some improbable plane crashes). Determining these causes might be computationally infeasible.

Unsustainable time delays. In many real-time situations, employing large sophisticated causal models to determine the most relevant causes may be infeasible. E.g., the sheer size of the causal model and the high frequency of events might make it impossible to compute the cause in a reasonable time frame. Moreover, non-linear dynamics such as chaos may also require too much time to detect and analyse. As a result, by the time relevant causes may be found, the system may have already changed, hence rendering this result irrelevant.

*2) **Finding Relevant Causes**:* Causes might be correct yet irrelevant. Relevance is relative and depends on the observer asking the question. Several key factors intervene here.

User Objective. *What is the purpose of finding a cause?* Finding 'the' cause of a conflict, or problem, may help to understand or to solve it.

If the aim is to *understand* the problem, then a relevant cause should indicate either: why the observation is wrong (e.g. a faulty sensor); or why the observation should have actually been expected (e.g. the sudden breeze shouldn't be surprising because the window is open). This may, in turn, lead to recursive questioning (e.g. why is the sensor failing? why is the window open?) – cf. the issue of 'infinite regression'.

If the aim is to *solve* a problem, then a relevant cause should help identify an effective action for achieving that. E.g. if a user asks 'why is it cold inside?', the answer 'because the temperature is $10°C$' is irrelevant, whereas 'because the window is open' is relevant as it helps fix the issue by closing the window. Still, the cause of a problem may be irrelevant for fixing it. E.g. when a patient's problem is a sprained ankle, one should first focus on *what to do* to fix it, rather than inquire *why* it happened. Such action-driven approach also applies when the cause cannot be found within given constraints.

Finding a cause that points to an effective action will also help select the search space scope and abstraction level (cf. subsec. III-A-1). E.g., once a home owner learns that an observed power outage is caused by a regional blackout, there is little point in searching for further causes – that would bring little useful information on what to do about it. Hence, the search scope may be limited to the corresponding action scope. Likewise, the granularity of available actions may constrain the abstraction levels relevant to cause searching (e.g. 'if your only tool is a hammer then every problem looks like a nail').

User knowledge and context. *Who is asking the question and what do they know?* As relevant cause(s) help to understand or fix conflicts, they depend on the observer's knowledge, expectations, beliefs and context. These aspects can help narrow the scope of the abduction process and guide it in suitable directions. E.g., considering the question 'why is it raining ?'. To a curious child, the relevant answer may be 'because water evaporated in the clouds where droplets formed and became too heavy to stand-up in the air'. To a sunshine-expecting meteorologist, a relevant answer may be 'because your predictive model missed a key parameter'. To a house inhabitant, the relevant cause may be that 'there is a hole in the roof'. Hence, a relevant cause depends on the nature of the conflict generating the question, which is user-dependent.

IV. Causality Research Roadmap for CAS

Common causal models such as SCM and CBN are, in principle, sufficiently generic to represent causal relations in any system, including CAS, or SISSY. Yet, we still lack reusable methods and sub-theories to adapt such generic models to the specific difficulties inherent in complex, dynamic and unpredictable systems. At the same time, dedicated research areas have been studying these specific difficulties quite extensively and provided a wide range of modeling and analysis tools.

Based on the challenges and shortcomings identified above, we point to several research areas that may be relevant for extending existing causality research. While most identified areas pertain to computer science and engineering, other CAS-related areas are also relevant here. E.g., research in modeling human-like knowledge representation and reasoning (e.g. cognitive science, language evolution and symbolic AI) provide promising research directions for addressing some of the challenges identified here. For instance, cognitive modeling has been employed to provide interpretable explanations [17] or to use inspiration from human cognition for better system design. Research in animal cognition and behaviour may also be relevant – e.g. evolution of movement and linguistic symbols [18] or self-extensible symbol systems [19].

We suggest investigating how causal modeling and analysis research may be extended by adopting and integrating relevant techniques from these adjacent areas. To this end, we propose a research road-map as a list of promising investigation directions for furthering current causality theory.

A. Towards Open Self-Adaptive Causal Modeling

Generally, the purpose of causal models is somewhat similar to that of system dynamics models. Both model types aim to provide insights into system behaviour – i.e. how do variable values change over time, in relation to each other? and how do certain variable values or inter-variable structural properties impact global behaviour? From this perspective, adopting the rich expressiveness and analysis tools available from dynamic systems modeling (e.g. differential equations and agent-based simulations) would bring critical complements to causal modeling and abduction techniques.

Models at runtime (M@R) [20], self-modeling and self-awareness [21]. Dynamic system modeling is key to causality representations, especially when models can be updated automatically, according to relevant changes in the targeted system and its environment.

Just-in-time, partial, elastic modeling. To enable causal models to scale to large systems and frequent changes, they should only consider modeling a limited system scope, in terms of domain and time length. Ideally, such partial models would be created at runtime, and be question-dependent [22]. The modeling scope may then be adapted dynamically – e.g. extending progressively over subsequent iterations.

Multi-paradigm models (MPM) [23]. MPM aims to represent systems from different, interrelated view points, via various formalisms and abstraction levels. This approach is essential for scaling causality models to large CAS.

Adaptive multi-scale models [7], or Lensing [10], [11]. Causal modeling should be able to represent a targeted system at different abstraction levels [3] and allow abduction processes to move across levels dynamically to find relevant causes [12], [11]. Causal blankets [24] and Markov blankets [25] provide promising directions to model the extent of individual systems and the interaction with others in complex environments. Causal models for systems with complex or chaotic dynamics may also need to represent feedbacks between abstraction levels [7].

Temporal and dynamic causality. Causal modeling should systematically include considerations from temporal causality [26] or Dynamic Causal Modeling [27], as well as extend its usage of temporal abstractions and unknown temporal scopes.

Dynamic discovery and integration of new variables, concepts or symbols [19]. Causal models should be able to discover and integrate new variables to account for dynamic updates in system observables (e.g. new sensors) or relevant abstractions (e.g. variables that are only observable indirectly via their effect). For instance, [28] aims to infer abstraction levels based on data-driven machine learning techniques.

Dealing with unknown causal effects. Problem-solving processes may consider alternative approaches when reliable causes cannot be found – e.g. trial-and-error heuristic methods, within safety constraints, may be better than doing nothing or doing the wrong thing based on a skewed model.

B. Towards Viable Abduction Processes for Relevant Causes

Enabling CAS to respond to changing conditions in the environment, including the interaction with other systems, may require identifying causes effectively and efficiently.

Multi-paradigm abduction methods. Different abduction methods may be best suited to different questions, context and targeted systems. For relatively simple or complicated systems, classic backtracking-driven abduction may suit. Yet, complex or chaotic systems may require alternative methods. When the basic system inner-workings are known, (partial) online simulations may be viable. In highly-dynamic contexts, this may imply generating short-term models on-the-fly (e.g. self-modeling, self-aware systems). Alternatively, when large datasets are available, machine learning techniques can be employed to correlate previous events to observed effects [22].

Self-diagnosis of causal model relevance. A dedicated function enabling systems to determine whether their current causal model is relevant in the present context would help trigger a model-update process, select a more suitable abstraction level or an appropriate abduction method; or decide that the system is ill-equipped and prefer to use resources elsewhere.

Self-diagnosis of the abduction process. To select a viable abduction process, or to determine whether abduction is even desirable in the context, one must first establish the nature of the system at hand. A question arises whether we can identify a system's particular non-linear dynamics (e.g. chaotic or complex) – manually or via a self-diagnosis process. E.g., in engineered systems, the very fact that the system is chaotic may represent a relevant cause for observed problems, as fixing these problems would imply restructuring the system.

Iterative, dialogue-based model adaptation and recursive abduction methods. To guide model adaptation and abduction processes, iterative approaches based on automated tests or user dialogue may be employed [17]. This would push the problem of cause detection within large search spaces onto a user or a self-diagnosis process; and help fix the the infinite cause regression problem.

V. Conclusions

This paper aimed to highlight several key outstanding challenges in applying common causal modeling and abduction techniques to CAS, including SISSY. We emphasised the fact that current mainstream methods, such as SCM and associated cause-searching processes (abduction) are sufficiently generic to apply to most systems, yet require significant effort to do so. This is due to particular characteristics of open, large CAS – e.g. size, non-linear dynamics, unpredictability, evolution – requiring a paradigm shift in the development, maintenance and employment of 'classic' causality-related techniques.

We acknowledged the fact that highly-relevant expertise for addressing these challenges is already available from related areas – e.g. complex (adaptive) system modeling and model-driven self-management; and specific application domains from engineering; cognitive science; social, economic and environment predictions, to only mention a few.

Once suitable methods are conceptualised , e.g. using expertise from these fields, they can be derived into reusable domain-specific tools applicable to real complex systems. This preliminary, non-exhaustive analysis provides an initial basis and generic research direction towards the necessary paradigm shift for rendering causality modeling and analysis tools more suitable for CAS applications.

References

[1] J. Pearl, *Causality*. Cambridge university press, 2009.

[2] J.-L. Dessalles, "A cognitive approach to relevant argument generation," in *Principles and Practice of Multi-Agent Systems*. Springer International Publishing, 2016, pp. 3–15.

[3] K. Fadiga, E. Houzé, A. Diaconescu, and J.-L. Dessalles, "To do or not to do: finding causal relations in smart homes," *Proceedings of the International Conference on Autonomic Computing and Self-Organizing Systems*, pp. 110–119, 2021.

[4] M. J. Vowels, N. C. Camgoz, and R. Bowden, "D'ya like dags? a survey on structure learning and causal discovery," *ACM Computing Surveys*, vol. 55, no. 4, pp. 1–36, 2022.

[5] R. Guo, L. Cheng, J. Li, P. R. Hahn, and H. Liu, "A survey of learning causality with data: Problems and methods," *ACM Computing Surveys (CSUR)*, vol. 53, no. 4, pp. 1–37, 2020.

[6] J. Y. Halpern, "A modification of the halpern-pearl definition of causality," in *Proceedings of the International Conference on Artificial Intelligence*, 2015, pp. 3022–3033.

[7] A. Diaconescu, L. J. Di Felice, and P. Mellodge, "Multi-scale feedbacks for large-scale coordination in self-systems," in *Proceedings of the International Conference on Self-Adaptive and Self-Organizing Systems*, 2019, pp. 137–142.

[8] A. Diaconescu, L.-J. Di Felice, and P. Mellodge, "An information-oriented view of multi-scale systems," in *IEEE International Conference on Autonomic Computing and Self-Organizing Systems Companion (ACSOS-C)*, 2021, pp. 154–159.

[9] P. Gardenfors, *Conceptual spaces: The geometry of thought*. MIT press, 2004.

[10] A. Diaconescu, D. W. King, K. L. Bellman, C. Landauer, and P. R. Nelson, "Lensing: It is all about perspective," *Proceedings of the International Conference on Autonomic Computing and Self-Organizing Systems Companion*, pp. 119–125, 2022.

[11] A. Diaconescu, D. King, K. Bellman, C. Landauer, and P. Nelson, "Towards systems that dynamically change and evaluate abstractions," in *Proceedings of the International Conference on Cognitive Situation Management*, 2023.

[12] A. Diaconescu, E. Houzé, J.-L. Dessalles, H. Vangheluwe, and R. Franceschini, "Multi-scale model-based explanations for cyber-physical systems: the urban traffic case," in *Proceedings of the International Conference on Model Driven Engineering Languages and Systems: Companion Proceedings*, ser. MODELS '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 684–691.

[13] J. C. Flack, "Coarse-graining as a downward causation mechanism," *Philopsophical Transactions Royal Society A*, vol. 375, no. 2109, 2017.

[14] C. M. Barnes, L. Esterle, and J. N. A. Brown, ""When you Believe in Things that you don't Understand": the Effect of Cross-Generational Habits on Self-Improving System Integration," in *Proceedings of the International Workshops on Foundations and Applications of Self* Systems*, 2019, pp. 28–31.

[15] S. Beckers, "Causal sufficiency and actual causation," *Journal of Philosophical Logic*, vol. 50, no. 6, pp. 1341–1374, 2021.

[16] S. Triantafyllou, A. Singla, and G. Radanovic, "Actual causality and responsibility attribution in decentralized partially observable markov decision processes," in *Proceedings of the Conference on AI, Ethics, and Society*, 2022, pp. 739–752.

[17] E. Houzé, A. Diaconescu, J.-L. Dessalles, and D. Menga, "A generic and modular reference architecture for self-explainable smart homes," in *2022 IEEE Intl Cnf on Autonomic Computing and Self-Organizing Systems (ACSOS)*, 2022, pp. 101–110.

[18] K. Bellman and L. Goldberg, "Common origin of linguistic and movement abilities," *The American journal of physiology*, vol. 246, pp. R915–21, 07 1984.

[19] C. Landauer and K. Bellman, "Symbol systems in constructed complex systems," in *Proceedings of the International Symposium on Intelligent Control Intelligent Systems and Semiotics*, 02 1999, pp. 191 – 197.

[20] N. Bencomo, R. France, B. H. Cheng, and U. Aßmann, *Modelsrun.time – Foundations, Applications, and Roadmaps*. Springer, 2014.

[21] K. L. Bellman, C. Landauer, P. Nelson, N. Bencomo, S. Götz, P. Lewis, and L. Esterle, "Self-modeling and self-awareness," in *Self-Aware Computing Systems*. Springer International Publishing, 2017, pp. 279–304.

[22] S. Reyd, A. Diaconescu, and J.-L. Dessalles, "CIRCE: a scalable methodology for causal explanations in cyber-physical systems," 5th IEEE International Conference on Autonomic Computing and Self-Organizing Systems (ACSOS), 2024, to be published.

[23] M. Amrani, D. Blouin, R. Heinrich, A. Rensink, H. Vangheluwe, and A. Wortmann, "Multi-paradigm modelling for cyber–physical systems: a descriptive framework," *Software and Systems Modeling*, vol. 20, 06 2021.

[24] F. E. Rosas, P. A. M. Mediano, M. Biehl, S. Chandaria, and D. Polani, "Causal blankets: Theory and algorithmic framework," 2020. [Online]. Available: https://arxiv.org/abs/2008.12568

[25] M. Kirchhoff, T. Parr, E. Palacios, K. Friston, and J. Kiverstein, "The markov blankets of life: autonomy, active inference and the free energy principle," *Journal of The Royal Society Interface*, vol. 15, no. 138, p. 20170792, 2018.

[26] C. K. Assaad, E. Devijver, and E. Gaussier, "Survey and evaluation of causal discovery methods for time series," *Journal of Artificial Intelligence Research*, vol. 73, pp. 767–819, 2022.

[27] P. K. Rubenstein, S. Bongers, B. Schölkopf, and J. M. Mooij, "From deterministic odes to dynamic structural causal models," *arXiv preprint arXiv:1608.08028*, 2016.

[28] M. Pol and A. Diaconescu, "A data-driven approach for modeling unknown multi-scale systems," in *2023 IEEE Int Conf on Autonomic Computing and Self-Organizing Systems Companion (ACSOS-C)*, 2023, pp. 35–40.