Causal AI-Powered Event Interpretation: A Cause-and-Effect Discovery for Indoor Thermal Comfort Measurements

Bukhoree Sahoh, Charlee Kaewrat, Kirttayoth Yeranee, Nichnan Kittiphattanabawon, and Mallika Kliangkhlao

Abstract—Unexplainable indoor thermal comfort events from black-box models influence people to distrust suggestions from decision-support systems and ask for help from engineers and practitioners that are labor intensive and time consuming. These problems come from unknown cause and effect in the environments that cause the system not to produce explainable outcomes. This study proposes the cause-and-effect discovery for indoor thermal comfort events that help systems make human-like explanations to overcome these issues. The research contributions consist of three essential points. The first is perceptions based on the Internet of Things (IoT) technology widely used in thermal comfort fields. Wu et al. [1] proposed predictive approaches for indoor thermal comfort conditions on time. The results showed that IoT and ML could reach excellent results but blind explanations [6]. When the outcomes were chosen and how historical data, but they could not produce reasons such outcomes were chosen and why not something else. These are current limitations of “black-box” ML, where they can not something else. These approaches are applied to real-world thermal comfort where occupants need reasonable details from systems rather than specific outcomes, it becomes a problem. Without reasonable explanations from the automation system for occupants, it causes them to distrust predicted outcomes, it becomes a problem.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/
outcomes and reject the automation system’s recommendation. Therefore, they end up asking engineers and technicians to diagnose physical site problems manually. Explainable AI (XAI) offers transparent technology that ML models can explicitly express how such outputs were predicted. It is a model for the decision-making process that can be seen by occupants who need explanations of how the outcomes are computed. Tree-based MLs are applied to recent XAI as glass-box models, such as decision trees, random forests, and gradient-boosted decision trees [7]. Sachan et al. [8] and Lundberg et al. [9] employed glass-box models in critical systems to produce logical details to support decision makers. They claimed that the models could reach high accuracy and good explainability, ready to apply to real-world applications. However, the tree-based MLs are the human-in-loop model, visualize knowledge based on graphical symbols so that humans can see how the output comes up, but they do not encode how nature works. In simple words, the system does not reflect real-world situations but returns statistical correlations from historical data that lack human-like interpretations. For instance, models may encode correlations between the overconsumption of electric current and hot outdoor. However, we understand that electric current and outdoor environment do not affect the change of each other, and we do not employ one to explain another but employ cause-and-effect concepts.

The cause-and-effect is a fundamental principle of human-like interpretations that lets systems imitate human reasoning to explain events of interest. The cause-and-effect-based systems are a native model (not necessarily human-in-loop) that can plausibly explain situations based on reality in a human-like intelligence manner. Suppose we are under an excellent indoor climate where the outdoor environment is terrible. In that case, the air condition system should be turned on to control indoor thermal comfort, and it must consume high electricity. People can make sense of the cause-and-effect explanations. However, current XAI based-systems cannot infer and produce human-like interpretations because they do not know what correlation does not imply causation.

This research proposes discovering cause-and-effect ML models for indoor thermal comfort understanding to overcome this problem. It is for causal AI-powered decision-making systems that can automatically infer and produce explanations in the manner of human-like intelligence. The main contributions of this research are as follows.

1) To propose the human-like perception based on IoT technologies using random variable systems to measure the environmental thermal comfort. This encodes the sense of engineers and technicians that detects fault events from the sensory signals.

2) To propose qualitative knowledge representation using the graph-based structural causal model (SCM) as the ground truth of how engineers and technicians explain thermal comfort events based on the cause and effect.

3) To prove that causal discovery algorithms can determine cause-and-effect ML from observational data converged to qualitative knowledge in the manner of human-like intelligence.

Thermal comfort decision-making systems, causal discovery, and related applications are examined in Section II. Material and methods are detailed in Section III. The overview system of thermal comfort understanding is proposed in Section IV. Section V concentrates on the environmental perception using IoT technologies, with Section VI focusing on transforming the perception to thermal comfort variables. Section VII performs experiments and analyses how cause-and-effect ML models imitate human understanding, and conclusions appear in Section VIII.

II. RELATED WORK

This section observes the recent trends for making decisions in indoor thermal comforts based on IoT and ML models, highlighting the decision-making process’s limitations.

A. Related Work Based on IoT and ML

Thermal comfort considers how to control dynamic factors that influence the occupants’ thermal sensation in an indoor environment. de Dear and Brager [10] introduced the occupant factors (e.g., clothing and activities), environment factors (e.g., temperature and humidity), and air condition factors (e.g., air temperature and airflow speed) that cause the thermal comfort events to change over time [11]. We employ this background knowledge to review related works for interpreting indoor environments.

IoT is an agent’s imitations of human-like perception to collect and share environmental signals. It lets software agents (e.g., decision-making systems) connect to the physical environment and monitor events automatically. Indoor thermal comfort employs an IoT technology to transform the analog world into a digital world that can minimize the need for human action. Wall et al. [12] and Morresi et al. [13] designed and developed an IoT-based agent for monitoring indoor air quality. They summarized that it is a helpful technology that benefits occupants and engineers control the indoor environment. Cao et al. [14] measured environmental factors for sleep air quality based on IoT. They found that IoT technologies could help agents automatically identify indoor thermal comfort and benefit human life quality.

The IoT-based research aimed to perceive the environment’s signals and determined events relevant to thermal comfort. The fundamental limitations are that they did not focus on event interpretation to explain the reasons. It causes systems could not reach the information of the human-like intelligent level. This suggests that IoT alone could not complete software agents to achieve human-like interpretation and needs additional technologies to fulfill the gap.

An ML is a mechanism of an agent’s brain to mimic human-like intelligence that can learn particular tasks and improve when gaining more experience. It lets software agents decide and cope with problems under uncertain situations [15]. The integration between ML and IoT technologies helps agents observe and analyze abnormalities in thermal comfort that can deal with issues early. They can reduce time processing and human action for solving complex problems. Ma et al. [16] and Somu et al. [17] introduced an ML-and-IoT-based agent.
to predict indoor thermal comfort aligned with occupant preferences. They claimed that the agent could help occupants adapt their behavior correctly even in an uncertain situation. Yang et al. [18] and Peng et al. [19] employed ML-and-IoT technologies to control indoor thermal comfort adapted to occupants’ behaviors. They discussed that the agent could handle indoor thermal comfort, benefiting occupants’ well-being and helping energy savings. Liu et al. [20], [21] applied tree-based MLs (e.g., decision tree and random forest) for indoor thermal comfort that allowed occupants to understand the way systems predict the outcomes based on if-path patterns.

The ML-based research focused on automatic monitoring and analysis agents that could predict outcomes based on data associations. The limitations are that they did not contribute to reasonable explanations based on causations. In other words, they did not concern about how to teach them to interpret and explain problems in the manner of human-like reasoning. Explainable information is essential because engineers must understand how nature works based on the cause and effect. They need to solve how and why agents generated such outcomes rather than summarizing based on statistical correlations.

B. Explainable Information in Decision-Making System

Explanations are the intellectual ability of human beings to communicate and exchange their experiences. People can introduce scientific explanations to understand how nature works [22]. A cause-and-effect encodes systematic understanding that helps software agents imitate human-like intelligence to manufacture cause-and-effect information.

Scientific explanations play a crucial role in indoor thermal comfort when engineers and practitioners must deal with critical problems like occupants calling for help since they feel uncomfortable. The goal of scientific explanations of indoor thermal comfort is to answer essential questions about how engineers interpret and explain indoor thermal comfort. The critical questions are advanced levels of human-like intelligence, and their answers need the knowledge to synthesize whether the question is Association, Intervention, or Counterfactual [23]. Different questions need different expertise to plan and provide potential solutions. The examples of critical questions are shown in Table I.

Table I shows that the first question can be asked when occupants suffer from discomfort. The engineers can answer such questions by observing how the events have co-occurred. It is simple information without assumptions of actions. The second is intervention; the engineers and practitioners must make an effort to answer this question. It is based on the potential outcome using their experiences and experiments—the intervention models the effect of the “What if we do?” ‘s activity based on actions that the result will be the future prior knowledge. The final is counterfactual, where they imagine going back in time and doing something different. It is unreal, but people always go for this question to interpret real-world events. The counterfactual is a high level of human thought that allows people to imagine things that do not exist in reality but may happen, called creative thinking.

{| **TABLE I**
| EXAMPLES OF CRITICAL QUESTIONS IN INDOOR THERMAL COMFORT |
|---|
| **No** | **QUESTION IN INDOOR THERMAL COMFORT** | **TYPE** |
| 1 | What does improper maintenance of the air conditioner tell us about discomfort? | Association |
| 2 | Suppose improper maintenance of the air conditioner causes the room to be uncomfortable. What if we maintain it? | Intervention |
| 3 | What would have been if we had maintained the air conditioner in the last two months? | Counterfactual |

The questions have happened every day for human beings, and they exchange knowledge through answering the questions. However, recent agents cannot do so because they cannot compute cause-and-effect and produce scientific explanations. The results are that people do not trust agents’ suggestions and reject cooperation between humans and agents. In this way, critical questions need new technologies to compute answers based on cause-and-effect concepts.

III. MATERIAL AND METHODS

This section presents potential material and methods based on the cause-and-effect to address current issues. They consist of an SCM, causal graph, and causal discovery.

A. Structural Causal Model and Causal Graph

SCMs encode cause-and-effect relationships between events in the environments where the cause event happens, influencing the effect event. SCM employs systems of the structural equation to encode cause-and-effect relationships, and it consists of three sets: 1) exogenous variable \( U = \{ U_1, U_2, \ldots, U_j \} \); 2) endogenous variable \( V = \{ V_1, V_2, V_3, \ldots, V_j \} \); and 3) function \( f = \{ f_1, f_2, f_3, \ldots, f_k \} \). Random variables model \( U \) and \( V \) and function (e.g., probabilistic density function) connect them based on the cause-and-effect [24]. For example, cause event \( C \) and effect event \( E \) are endogenous variables that we can observe. We recognize that \( C \) directly influences \( E \) to occur. Therefore, \( f_c \) and \( f_e \) semantically provide the chances of both events using assignment operator \( \leftarrow \), and the equation is

\[
C_i \leftarrow f_c(U_i).
\]  

\( U_c \), an exogenous variable, represents unknown factors in the universe that we cannot measure, which causes \( C_i \) to happen. \( C_i \) is the event that we can observe in the environment, \( U_c \) assigns the chance to the \( C_i \) through function \( f_c \), representing the cause-and-effect operator between \( U_c \) and \( C_i \). Therefore, \( C_i \) assigns the chance to the effect event \( E_i \), and the structural equation of the cause-and-effect between them is

\[
E_i \leftarrow f_e(C_i, U_c).
\]
Fig. 1. Cause-and-effect graph modeling relationships between observable variables, $C_i$ and $E_i$, where $C_i$ directly causes the value of $E_i$. The solid-directed line represents cause-and-effect relationships between them. $U_c$ and $U_e$ are unobservable variables in the universe but cause $C_i$ and $E_i$ to occur. The dashed-directed line represents cause-and-effect relationships between unobservable and observable variables.

$f_c$ and $f_e$ encode relationships between cause-and-effect with their uncertainty in the form of probabilistic density.

In conclusion, SCM represents the semantic dependencies of $C_i$ directly causes $E_i$ using (1) and (2) based on the cause-and-effect operator. Instead of equations, it can be written to graphs that fully convey its semantics.

Causal graphs simplify the understanding of nature to nodes and edges where nodes model $U$ and $V$, and edges model $f$ between them. They visualize structural knowledge from data-driven and prior knowledge that human and software agents can understand and interpret the meaning. The cause-and-effect graph between $C_i$ and $E_i$ is shown in Fig. 1. Fig. 1 graphically models the cause-and-effect of SCM (1) and (2) between $C_i$ and $E_i$, $U_c$ and $U_e$ and their edges can be omitted in the network, making it more straightforward. The graph represents how $C_i$ directly affects $E_i$ and not the other way around.

The functions of the cause-and-effect model can be expressed in conditional probability according to the following:

$$P(X_1, X_2, X_3, \ldots, X_n) = \prod_{i=1}^{n} (P(X_i | Pa(X_i))) \quad (3)$$

where $X_i$ stands for a set of the effect variables $E_i$, and $Pa(X_i)$ stands for cause variables $C_i$. The cause-and-effect model can transfer information between them using Bayes’ theorem using the following:

$$P(\text{Effect} \mid \text{Cause}) = \frac{P(\text{Cause} \mid \text{Effect})P \times P(\text{Effect})}{P(\text{Cause})} \quad (4)$$

Bayes’ theorem lets the model compute how the cause-and-effect variables interchange information using basic mathematics solutions (e.g., multiplied by a $P(\text{Cause})$ and divided by $P(\text{Cause} \mid \text{Effect})$ on both sides).

It is highly potential to apply the cause-and-effect graph to model indoor thermal comfort. For example, we understand that the status (on or off) of an air conditioner $(A_C)$ can control the comfort of indoor climate $(I_C)$, and it can be represented by the causal graph $A_C \rightarrow I_C$ (mostly omitted $U_{ac}$ and $U_{ic}$). On the contrary, we understand that events in indoor climates do not affect the air conditioner. In this way, cause-and-effect graphs help software agents compute semantics between events and infer how such events happen and why it is not the other way around. It imitates the manner of human-like intelligence to reason and explain events plausibly. AI-powered decision-making systems should employ them as mechanisms to infer and interpret the reasons behind their predicted outcomes to support engineers, practitioners, and occupants.

B. Causal Discovery

Causal discovery is a scientific exploration of the cause-and-effect structures between events in systems’ random variables given observational data [25]. It models the cause-and-effect using graphs represented by nodes and edges. Nodes encode random variables, and edges encode relationships between them where cause nodes point to effect nodes. In other words, causal discovery sketches the cause-and-effect relationships in the universe by inferring the embedded meaning underlies observational data in the manner of human-like interpretations. The causal discovery assumes that universal systems (complete cause-and-effect graph) produce observational data with particular reasons or causes [26].

Discovering the cause-and-effect structures is complicated when observational data has multiple types, fast productions, and a considerable volume. Determining by domain experts, such as engineers and practitioners is labor intensive and time consuming. Therefore, automatic learning algorithms challenge imitating human-like abilities to understand the system variables and define knowledge within observational data [27]. Causal discovery algorithms can be categorized into two standard types: 1) constraint based and 2) score based [28].

Constraint-based algorithms acquire cause-and-effect graphs based on statistical significance tests using conditional (in)dependence and $d$-separation [29]. They begin with fully connected edges between random variables and compute cause-and-effect relationships based on the Markov equivalent class. It implies that different statistical significances based on conditional testing help $d$-separation expose cause-and-effect structures given observational data. They delete non-cause-and-effect relations based on $d$-separation if there are no statistical significances between random variables and causally connect if they do. In contrast, score-based algorithms learn the cause-and-effect graphs by measuring the best fit causal model (the highest score) given observational data. They initiate empty-relationship graphs and compute probabilistic scores based on the Bayesian information criterion (BIC) [30]. BIC computes the scores of likelihoods from observational data, given cause-and-effect graphs, where the graphs may vary by adding and removing cause-and-effect directions.

In conclusion, causal discovery algorithms automatically define cause-and-effect structures between random variables in thermal comfort understanding systems given sufficiently observational data.

IV. OVERVIEW SYSTEM OF THERMAL COMFORT UNDERSTANDING

Causal AI-powered decision-making systems for indoor thermal comfort understanding need a systematic approach to process environmental evidence and interpret its meaning to support decision makers. Morresi et al. [13] and
Somu et al. [17] introduced outstanding indoor thermal comfort architectures for automatic systems based on prediction frameworks. This article extends their basic concept and proposes the overview architecture to discover indoor thermal comfort. The architecture is enhanced by adding a new cause-and-effect modeling feature as a fundamental element to achieve human-like intelligence for event interpretation. Our proposed system architecture is shown in Fig. 2.

Fig. 2 shows that the architecture consists of two parts: 1) the environment where indoor thermal comfort events are happening and 2) the computing platform that transforms and interprets the event semantics for supporting engineers, practitioners, and occupants to understand situations and make decisions to respond appropriately.

The environment is a perception of phenomena that occupants should be aware of its changes. It has two sides: 1) indoor (the white space) and 2) outdoor (the light-brown area), and both are highly correlated to each other. Understanding environmental factors help engineers and practitioners take action on time. We apply IoT technologies to encode environmental factors as human-like perceptions relevant to humidity, temperature, air conditioner status, and electric energy consumption.

The computing platform is human-like thinking that explains events of interest based on environmental perception as system inputs. It produces cause-and-effect explanations of how and why such events are happening. We utilize principal data analysis and cause-and-effect ML to construct the human-like-brain computing that consists of two subprocesses: 1) online process components represented in a solid line and 2) offline process components represented in a dashed line. Offline process components produce the ready-made cause-and-effect model that online process components can interpret real-world events and represent the outcome using human-understandable visualization.

The highlight of this architecture is the event interpretation component that differs from a traditional architecture. It helps causal AI-powered decision-making systems interpret why and how events have occurred. The following section will develop how the event interpretation in the architecture can be implemented in indoor thermal comfort.

V. ENVIRONMENTAL PERCEPTION

The IoT-driven approach plays a critical role in decision-making systems for indoor thermal comfort. It represents human-like perceptions to sense how the environmental factors are changing. Understanding indoor thermal comfort should consider the cycle of comfort control system consisting of three components: 1) perception of the environmental factors (Environment) helps humans deal with 2) their comfortable feeling (Human Comfort) and may decide to utilize 3) such air conditioner (HVAC System) to control the environment if they are uncomfortable [31].

We design and develop the environmental perception component based on the thermal comfort understanding cycle using two processes. The first is IoT-based sensors for reading analog signals. The second is digitization based on analog-to-digital transformations that are thermal comfort factor representations aligned with system architecture in Fig. 2. Sensor boxes are set into two positions in indoor and outdoor environments: 1) the first measures indoor temperature and humidity and 2) the second measures outdoor temperature and humidity—moreover, the second measures electric energy consumption in cases of occupants turning on HVAC systems. The design and development of environmental measurement factors are described in Table II.

Table II represents factors relevant to cycle control components: possible ranges, sensor models, and error rates.

### TABLE II

| FACTOR                  | RANGE   | SENSOR MODEL          | SENSOR ERROR | CYCLE CONTROL COMPONENT |
|-------------------------|---------|-----------------------|--------------|-------------------------|
| Indoor Temperature      | 0-50 °C | DHT22 AM2302          | ± 0.2°C      | Human Comfort           |
| Indoor Humidity         | 0-100% RH| DHT22 AM2302          | ± 1.0% RH    | Human Comfort           |
| Outdoor Temperature     | 0-50 °C | AM2315 I2C Single Bus | ± 0.3°C      | Environment             |
| Outdoor Humidity        | 0-100% RH| AM2315 I2C Single Bus | ± 2.0% RH    | Environment             |
| Electricity             | 0.00-20.00 Amp | SCT-013               | ± 0.2 Amp    | HVAC System             |
DHT22 model reads analog signals of indoor temperature and humidity factors representing human comfort. The AM2315 model is a waterproof sensor that browses analog signals of outdoor temperature and humidity factors representing the environment. An SCT-013 model scans analog electricity usage signals that are considered an indirect measurement of the HVAC system behaviors. The measurements of the right factors are critical issues that are needed to evaluate. We chose them based on the studies of thermal comfort measurements [32], which help us use fewer sensor models but still work effectively.

We set up a box set to hold sensors for data streaming suitable for installation in the occupant buildings, following the complete study proposed by Sahoh et al. [33]. Box sets are introduced based on cause-and-effect perceptions shown in Fig. 3.

Fig. 3(a) shows that the box set is compact and ready to observe the outdoor environment. Its size is measured in $10 \times 7.5 \times 2.5$ cubic centimeters (length × width × height): component 1) percepts temperature and humidity, 2) percepts electricity usage, 3) stores data backup (in case of Internet down), 4) voltage regulators handle fluctuations and supply voltages stably. The sensors are controlled by 5) a microcontroller based on ESP32 with a robust design and low-power consumption. In conclusion, our box sets are well-designed for occupants’ easy use and are durable for infrastructure problems, such as the Internet connection and voltage noise. We will collect the input signal and feed it into the causal computing platform.

Fig. 3(b) shows the two boxes installed sensing signals based on cause-and-effect. Indoor perception 1) collects effect signals built in the indoor environment streaming temperature and humidity. It was set close to the air conditioner unit since it can sense indoor factors like the air conditioner’s thermostat. Outdoor perception 2) collects cause signals from the outdoor environment and the air conditioner’s power consumption. It was placed outside the buildings that can perceive outdoor temperature and humidity. Moreover, it must connect to the electricity usage from the air compressor, measuring the HVAC system.

The following section will develop how to model such input based on human-like perceptions.

VI. CAUSAL MACHINE LEARNING

This development of causal ML for indoor thermal comforts is based on four perspectives: 1) design of thermal comfort factors; 2) design and development of the cause-and-effect model blueprint; 3) parameter estimations for the cause-and-effect model blueprint; and 4) causal structure discovery based on observational data.

A. Design of Thermal Comfort Factors

The thermal comfort situations have actual events, but their chance randomly happens. For example, we understand that indoor thermal comfort outcomes might be good, fair, harmful, or worst but do not exactly know which events will occur first. They may be caused by other events or unmeasured factors in the universe and need technologies to quantify and represent in machine-computable formats. Fortunately, random variable systems can encode the challenge of thermal comfort events using measurable functions and describe them into probability
distributions. They illustrate how events are likely to happen and allow humans and software agents to ask the questions as follows.

1) What is the likelihood of harmful indoor thermal comfort occurring in the afternoon?
2) If we set the air conditioner at 25 °C, what is the chance of harmful indoor thermal comfort in the afternoon?

Random variables mathematically describe unknown events and produce the questions’ answers under uncertain situations that help humans and software agents recognize the indoor thermal comforts in the same manner. We design the random variables aligned with the needs of measurement factors for indoor thermal comforts from Table II as shown in Table III.

Table III shows that five random variables and their states were discretized according to human-like understanding. It confirms that all thermal comfort signals are relevant to human-like experiences. T’s events are classified into standard periods encoding the meaning of time that can affect thermal comfort, such as the comfort in the afternoon must differ from the night. According to engineers and practitioners, H and P’s events were modeled on interpreting air conditioner status based on the specifications. They can determine whether air condition systems are working or not. E and I modeled the indoor thermal comforts and environments that employ background knowledge (according to references). They play a crucial role in explaining indoor and outdoor situations that humans use to feel their comfort level and decide how to adapt if the problems are worst.

The five random variables represent the models of how humans perceive signals and randomly encode semantic information so that they can understand and exchange such information and let software agents imitate human-like understanding.

Automatic interpreting of the events of thermal comfort needs deeper information to reach the human-like agreement of cause-and-effect between random variables. In other words, the indoors does not occur randomly but depend on the others. For example, we understand that indoor comfort may be worst if the HVAC system is off. However, if the outdoor situation has suitable temperature and humidity, they influence the indoors to be comfortable. This understanding needs human-like technologies to discover causations between random variables and explain why such events occur, not something else.

### B. Design and Development of the Cause-and-Effect Model Blueprint

This research conducts a human-like model to interpret how and why indoor thermal comforts occur. Therefore, we begin the design and development of the cause-and-effect model by interviewing and discussing this concern with three thermal comfort engineers who understand the thermal comfort systems (random variables). Four practitioners help us identify the events in the fields of thermal comfort (random variables’ states). Their experiences and knowledge can determine the causal mechanisms behind observational data generations.

We cooperated with engineers in academic fields (two thermal comfort engineers and one heat transfer engineer) to draw the causal structures based on correlation that does not imply causation, meaning random variables must causally relate to others. We collaborated with practitioners to identify unexpected events when intervening in how variable influences each other in the practical fields. We figured out the causal diagram based on random variables shown in Fig. 4.

Fig. 4 shows the cause-and-effect graph of indoor thermal comforts where I is a common effect of E and H that determine the structure of the collider. H is a common cause of I and P that name the fork’s structure. E sequentially connects T, and I called the chain structure. The cause-and-effect structures encode mechanisms behind data-generating models. The proofs of concepts of cause-and-effect structures, collider, fork, and chain, were proposed by Barber [36].

We can interpret that E and H (their parents or children) are highly correlated (change statuses together), but they do not exchange any information unless I has been conditioned. Suppose we want to control I to be comfortable in the summer’s afternoon where E is high temperature and humidity. In that case, we have to turn H on, and it may overwork since I situations entirely converge on E. It causes P to consume more energy to run H and forces I to be satisfied. T indirectly influences I through E, meaning different periods cause the outdoor environment to be in other states. In contrast, if H has a problem, it may consume P as before but cannot control I to be comfortable.

The cause-and-effect mechanism-based model has qualitative knowledge. It needs to quantify the qualitative experience into probability distributions and estimate their parameters using observational data.
method. It can estimate the parameters optimally even though the observational data are scarce and imperfect. It aims to infer the behavior of random variables given data rather than summarize its fixed quantity [37]. We employed a generalized Bernoulli distribution to encode the discrete events in our cause-and-effect model. The examples of \( P \)'s and \( H \)'s optimized parameters, which are the marginal distributions and conditional probability distributions between them, are shown in Fig. 5.

Fig. 5 exemplifies the parameters of \( H \) and \( P \) and their cause-and-effect relationships in the form of probability tables. For example, Fig. 5(a) shows the nature of the HVAC system works that system failure rarely occurs in the real world (8.06%) while most \( H \)'s system is off (37.75%). Engineers confirm that they are credible because 40% of the HVAC system is off during nonworking time. They claimed that \( H \)'s system overworking is approximately 19.07% acceptable because staff in the office building typically take time to recognize if the \( H \) has problems. Fig. 5(b) shows that most \( P \)'s behavior (if we do not consider the system off) is an average load (26.86%) because most office buildings’ \( H \) is properly working if they are regularly maintained. Overconsumption rarely occurs (9.51%) but is possible if organizations have no policy to support the \( H \)'s maintenance.

Fig. 5(c) encodes \( P \)'s knowledge conditioned on \( H \), \( P( P \mid H) \). We understand that \( H \) influences \( P \), which means we can approximate what \( P \) will be if we recognize the \( H \) information. For example, if \( H \) is turned off, \( P \) should be no power consumption (93.33%). In contrast, if \( H \) is working, \( P \) is expected to be average (53.07%) in general office buildings.

Engineers and practitioners agree with the proposed cause-and-effect model and confirm that distributions of \( P(H), P(P) \), and \( P( P \mid H) \) are intuitive and informative because they represent the conception of indoor thermal comfort that they used to support decision making.

In conclusion, Fig. 4 encodes human-like intelligent intuitions to understand the mechanisms behind indoor thermal comfort. Fig. 5 represents the cause-and-effect modeling fitted by observational data. In simple words, cause-and-effect models encode human-like intelligence, same as engineers and practitioners who use such knowledge to interpret indoor thermal comforts. However, a problem in this domain is the lack of practitioners who understand indoor thermal comfort. A significant research gap needs to be fulfilled, and causal structure discovery is a high potential for this concern.

### D. Causal Structure Discovery

This section aims to find the cause-and-effect models using algorithms, given observational data. The essential idea is that the system seeks to imitate the human-like intelligence of making hypotheses between random variables. For example, \( P \) may cause \( I \), \( I \) may cause \( P \), or there is a hidden variable connecting \( P \) and \( I \), like \( H \) may be a common cause between them. It lets software agents scientifically ask which variable causes another variable.

The learning process of a causal structure discovery assumes that observational data generated by environmental perceptions...
are ideally complete. Our model blueprint verifies the data generating process that is a source of our observational data sampling in indoor thermal comfort. It can be employed in a simulation approach to sampling the complete observational data for the learning process.

To make the discovery process fairly compared to expert knowledge, we generated complete observational data of 1,000,000 transactions from probability distributions based on our proposed blueprint model (see Section VI-B).

Therefore, we employed two general learning algorithms of cause-and-effect structure discovery that are: 1) Peter–Clark (PC) and FCI based on the constraint-based algorithm and 2) Hill-climbing algorithm based on log-likelihood information-theoretic scoring, exhaustive search algorithm based on Bayesian–Dirichlet equivalent uniform (BDeu), and greedy equivalence search algorithm based on Markov equivalence class based on the score-based algorithm.

We trained the algorithms using the transactions for discovering cause-and-effect relations, and the models are shown in Fig. 6.

Fig. 6 illustrates that all algorithms could correctly discover the I as a collider caused by E and H. However, the algorithms produce different types of directions that represent different meanings. PC algorithm manufactured undirected movement between T — E and P — H. The “—” means an unclear cause-and-effect relationship between two variables. In other words, the set of the graph from observational data can produce both T → E and T ← E. FCI algorithm generates “—” where “—” means that T can be the cause or effect of the others. For example, T — E can be translated to different structures: T → E, T ← E, or T ↔ E since both PC and FCI algorithms assume that cause-and-effect models discovered from observational data have no hidden variables or latent confounders [28].

GES discovered the cause-and-effect directions based on no unmeasured common causes in the observational data. It may generate the bidirected edges called partial ancestral graph (PAG) [30]. It is the reason why, for instance, T — E was produced in Fig. 6(c) that could be interpreted that both hold cause-and-effect relationships but the unknown direction, or there is a hidden variable (e.g., common cause) that affected both of them.

In contrast, Hill-climbing algorithms produced cause-and-effect graphs based on single directions. For example, Information-Theoretic Scoring-based Hill climbing in Fig. 6(b) discovers T → E, and BDeu-based Hill climbing in Fig. 6(d) discovers P ← H agrees with the model blueprint. They employ score functions to find the best cause-and-effect trends by adding or removing them based on the single-path assumption. They locally search for an optimal path based on the increased score functions and continuously iterates the search until there is no improvement.

However, there are incorrect cause-and-effect relationships compared with the model blueprint (e.g., opposite directions, bi directions, and irrelevant directions), and each algorithm produced different error directions. For example, Fig. 6(b) shows that the Hill-Climbing search algorithm based on log-likelihood information-theoretic scoring generated the
edge where $E$ directly affects $H$ that does not causally represent human-like understanding. Moreover, PC, FCI, and GES produce unmeasured edges (e.g., no arrowhead and bi arrowhead) that cannot indicate cause-and-effect directions and might cause agents to be confused in explaining how and why the thermal comfort events occur. In the next section, we will evaluate the cause-and-effect model performance of each algorithm to show how they converge to human-like intelligence.

### VII. Experimental Setup

The intelligent interpretations of software agents depend on the correctness of the cause-and-effect relationships between random variables. Our system evaluation measures the effectiveness of the cause-and-effect relationships produced by causal discovery algorithms.

#### A. Experimental Objectives

The objective is to measure the correctness of cause-and-effect directions between random variables since the agent intelligence depends on cause-and-effect identification. The measurement is based on the model blueprint as the ground truth approved by engineers and practitioners, which relationships against the model blueprint are considered fault relationships.

#### B. Model Testing Matrices

The evaluation matrices employ the fundamental principle of ML testing: precision, recall, and $F$-measure to measure the cause-and-effect directions in models. Precision represents the proportion of correct discovered cause-and-effect relationships against the total discovered cause-and-effect relationships. The recall represents the proportion of correct discovered cause-and-effect relationships against all cause-and-effect relationships. The $F$-measure represents an overall between precision and recall. Our measurement metrics can be defined as precision $= (TP/TP + FP)$, recall $= (TP/TP + FN)$, and $F$-measure $= (2 \times (Precision \times Recall)) / (Precision + Recall)$. Positive (TP) is correct cause-and-effect relationship between random variables that were discovered, false positive (FP) is an incorrect cause-and-effect relationship between random variables that were discovered, and false negative (FN) is the cause-and-effect relationship between random variables that were not discovered.

| ALGORITHM              | PRECISION | RECALL | $F$-MEASURE |
|------------------------|-----------|--------|-------------|
| PC                     | 0.20      | 0.33   | 0.25        |
| Log-Likelihood-based Hill climbing | 0.80 | 1.00 | 0.89 |
| BDeu-based Hill climbing | 0.75 | 0.75 | 0.75 |
| FCI                    | 0.33      | 0.50   | 0.40        |
| GES                    | 0.33      | 0.50   | 0.40        |

### C. Result Descriptions

According to the model blueprint, we consider that it discovered cause-and-effect relationships between random variables (e.g., Time $T$, Environment $E$, Indoor Comfort $I$, HVAC system $H$, and Power Consumption $P$). TP is a cause-and-effect relationship that is the directions similar to the model blueprint. FP is cause-and-effect relationships against the model blueprint. The missing cause-and-effect relationships are an FN. We simplify the conception by representing the TP, TF, and FN based on graphical models, as shown in Fig. 7.

Fig. 7 shows that the lines between random variables consist of black-solid, red-dashed, and red-solid. The black-solid line represents TP that cause-and-effect relationships in causal discovery agreed with the model blueprint. The red-dashed line represents FP as additional edges that did not exist in the model blueprint. The red-solid line represents the cause-and-effect relationships that causal discovery must produce but not. Undirected lines are considered two possible directions; for example, $T$ and $E$ in Fig. 7(a) have undirected line means $T$ causes $E$, and $E$ causes $T$.

The evaluation considered the ability of cause-and-effect direction discovery using our model testing matrices. We evaluated the cause-and-effect directions of five algorithms, and the results are shown in Table IV.

Table IV shows that PC, FCI, and GES could reach an unacceptable performance ($F$-measure is lower than 50%) since they discovered bidirectional edges based on the assumptions of no unmeasured confounders causing false-positive results. Their discovered false-positive edges are similar to the model blueprint but only miss edge directions. In this way, we can interpret that the confusions based on bidirectional edges are close to human-like abilities when they cannot identify what variable causes another. It means they have advantage gaps to be improved by applying new techniques to PC, FCI, and GES to handle the problems.

Hill-climbing algorithms discovered acceptable performance, especially the log-likelihood-based technique could reach 89% of the $F$-measure while the BDeu-based technique got 75%. However, when we considered the details, the log-likelihood-based technique discovered $E \rightarrow H$, an unreasonable direction since the environment did not causally influence the HVAC system information. In other words, the likelihood-based technique produces unnatural cause-and-effect relationships and suggests that the BDeu-based technique may challenge adjusting the learning process.

### D. Discussion

The overall structures of causal discovery models are similar to the model blueprint. The algorithms could discover $I$ as the collider that is a common effect controlled by $E$ and $H$. Collider structure is a proof of deep understanding when people describe two variables independently influence the third variable. The exciting point of collider $I$ is that it can separate $E$ and $P$ (including their parent and child) to be independent. For example, the change of $E$ and $T$ does not affect $H$ and $P$, which is a human-like manner of interpreting the real-world environment. In simple words, the shift in period does not...
influence electricity consumption unless \( I \) is controlled to be comfortable.

Cause-and-effect models are ground truth to infer relevant information in the domain of indoor thermal comfort. It did not fix with particular inputs and outcomes but allowed software agents and occupants to inquire about all possible states of random system variables. It will enable them to arbitrarily inquire information (e.g., all conditions of random variables) rather than compute input and produce predicted outcomes of a single task. When occupants ask about indoor thermal comfort information, cause-and-effect model-based software agents can causally answer the questions. For example, software agents observe \( I \) and \( P \), and may automatically offer \( H \)’s status on how it tends to be occurring. Moreover, the occupants may ask interventional questions, 1) “what is the most likely \( I \)’s status if \( T \)’s status is afternoon and \( H \)’s status is off?” or 2) “what is the most likely \( P \)’s status if current \( H \) is overwork.”

The same model can perform different tasks to answer the various questions based on SCM (see Section II-B), the computational core of causal inference [26]. Such questions are common for occupants to ask their engineers, and the cause-and-effect model lets software agents explain events in the manner of human-like intelligence. It confirms that the cause-and-effect model possesses human-like intelligence for indoor thermal comfort understanding.

However, PC, FCI, and GES generated low performances of cause-and-effect models since confusions of bidirection occur between two variables. The BDeu-based model has more minor disorders than PC, FCI, and GES; its cause-and-effect relationships are closer to the blueprint’s scope. Unlike log-likelihood-based hill-climbing, even though it got the highest performance, the cause-and-effect relationships are opposite to the model blueprint, which means the interpretations could be against human understanding if applied. This suggests that the improvements of current PC, FCI, GES, and BDeu-based models are challenging for future research. For example, we may allow expertise to cooperate in the design and development of the cause-and-effect discovery to adjust some misconceptions of algorithms [38].

VIII. CONCLUSION

This research has proposed a system process of cause-and-effect discovery for indoor thermal comfort measurements. It is to reason the observational data and manufacture explanations of indoor thermal comfort events that help occupants trust the system’s outcome. The system employs scientific perceptions to collect observational data on IoT technologies that imitate human sensation organs. Random variables and graphs are used to encode observational data into a machine-human readable format. We apply causal discovery algorithms to automatically determine cause-and-effect relationships in ML models from observational data that imitates human-like intelligence in the manner of engineers and practitioners.

Using accuracy, we evaluated cause-and-effect relationships in ML models and compared them against the model blueprint that engineers and practitioners determined. The results showed that ML models could discover cause-and-effect relationships close to the model blueprint given observational data. The \( F \)-measure of the BDeu-based model could reach 75%, and the log-likelihood-based model could get 89%. These results suggest that cause-and-effect ML models can...
produce reasonable explanations that help occupants trust such information and utilize it to make decisions.

In the future, we will set our experiment in different environments, such as complex HVAC systems in convention centers and department stores. The period will extend for all seasons in other provincial parts. These help the algorithms discover hidden knowledge that needs thermal comfort explanations.

We will adjust the algorithms based on cause-and-effect odd ratios to handle the undirected edges to determine directions. We will also consider Markov Chain Monte Carlo (MCMC) to sample data that lets algorithms identify independence between random variables in case of rare events.

REFERENCES

[1] R. Elnaklah, I. Walker, and S. Natarajan, “Moving to a green building: Indoor environment quality, thermal comfort and health,” Build. Environ., vol. 191, Mar. 2021, Art. no. 107592.

[2] Z. Wu et al., “Using an ensemble machine learning methodology-bagging to predict occupants’ thermal comfort in buildings,” Energy Build., vol. 173, pp. 117–127, Aug. 2018.

[3] Z. Wang, H. Onodera, and R. Matsuhashi, “Proposal of relative thermal sensation: Another dimension of thermal comfort and its investigation,” IEEE Access, vol. 9, pp. 36266–36281, 2021.

[4] C. Xu, H. Chen, J. Wang, Y. Guo, and Y. Yuan, “Improving prediction performance for indoor temperature in public buildings based on a novel deep learning method,” Build. Environ., vol. 148, pp. 128–135, Jan. 2019.

[5] N. A. Eltressy et al., “RF energy harvesting IoT system for museum ambiance control with deep learning,” Sensors, vol. 19, no. 20, p. 4465, 2019.

[6] C. Rudin, “Stop explaining black box machine learning models for high-stakes decisions and use interpretable models instead,” Nat. Mach. Intell., vol. 1, no. 5, pp. 206–215, May 2019.

[7] A. Rai, “Explainable AI: From black box to glass box,” J. Acad. Mark. Sci., vol. 48, no. 1, pp. 137–141, Dec. 2019.

[8] S. Sachan, J. B. Yang, D. L. Xu, D. E. Benavides, and Y. Li, “An explainable AI decision-support-system to automate loan underwriting,” Expert Syst. Appl., vol. 144, Apr. 2020, Art. no. 113100.

[9] S. M. Lundberg et al., “From local explanations to global understanding with explainable AI for trees,” Nat. Mach. Intell., vol. 2, no. 1, pp. 56–67, Jan. 2020.

[10] R. I. de Dear and G. S. Brager, “Developing an adaptive model of thermal comfort and preference,” ASHRAE Trans., vol. 104, pp. 145–167, Feb. 1998.

[11] A. Ghahramani, P. Galicia, D. Lehrer, Z. Varghese, Z. Wang, and Y. Pant, “Artificial intelligence for efficient thermal comfort systems: Recent applications, current limitations and future directions,” Front. Built Environ., vol. 6, p. 49, Apr. 2020.

[12] D. Wall, P. McCullagh, I. Cleland, and R. Bond, “Development of an Internet of Things solution to monitor and analyse indoor air quality,” Internet Things, vol. 14, Jun. 2021, Art. no. 100392.

[13] N. Montresi et al., “Sensing physiological and environmental quantities to measure human thermal comfort through machine learning techniques,” IEEE Sens. J., vol. 21, no. 10, pp. 12322–12337, May 2021.

[14] T. Cao, Z. Lian, S. Ma, and J. Bao, “Thermal comfort and sleep quality under temperature, relative humidity and illuminance in sleep environment,” J. Environ. Eng., vol. 43, Nov. 2021, Art. no. 102575.

[15] D. S. Vijayan, A. L. Rose, S. Arvindan, J. Revathy, and C. Amuthadevi, “Automation systems in smart buildings: A review,” J. Ambient Intell. Humaniz. Comput., vol. 1, pp. 1–13, Nov. 2020.

[16] N. Ma, L. Chen, J. Hu, P. Perdikaris, and W. W. Graham, “Adaptive behavior and different thermal experiences of real people: A Bayesian neural network approach to thermal preference prediction and classification,” Build. Environ., vol. 198, Jul. 2021, Art. no. 107875.

[17] N. Somu, M. R. G. Raman, and K. Ramamurthi, “A deep learning framework for building energy consumption forecast,” Renew. Sustain. Energy Rev., vol. 137, Mar. 2021, Art. no. 110591.

[18] C. Yang, B. Gunay, Z. Shi, and W. Shen, “Machine learning-based prognostics for central heating and cooling plant equipment health monitoring,” IEEE Trans. Autom. Sci. Eng., vol. 18, no. 1, pp. 346–355, Jan. 2021.

[19] Y. Peng, A. Rysanek, Z. Nagy, and A. Schlieter, “Using machine learning techniques for occupancy-prediction-based cooling control in office buildings,” Appl. Energy, vol. 211, pp. 1343–1358, Feb. 2018.

[20] Y. Liu et al., “Thermal preference prediction based on occupants’ adaptive behavior in indoor environments—A study of an air-conditioned multi-occupancy office in China,” Build. Environ., vol. 206, Dec. 2021, Art. no. 108355.

[21] Y. Liu, H. Chen, L. Zhang, and Z. Feng, “Enhancing building energy efficiency using a random forest model: A hybrid prediction approach,” Energy Rep., vol. 7, pp. 5003–5012, Nov. 2021.

[22] B. Sahoh and A. Choksuriwong, “Beyond deep event prediction: Deep event understanding based on explainable artificial intelligence,” in Interpretable Artificial Intelligence: A Perspective of Granular Computing, Studies in Computational Intelligence, vol. 937, W. Pedrycz and S. M. Chen, Eds. Cham, Switzerland: Springer, 2021, pp. 91–117.

[23] J. Pearl and D. Mackenzie, The Book of Why: The New Science of Cause and Effect. New York, NY, USA: Basic Books, 2018, p. 402.

[24] J. Kolner, P. W. Dalrymple, and R. Christiansen, Causal Inference in Statistics: A Primer, vol. 61. Chichester, U.K.: Wiley, 1986.

[25] F. Eberhardt, “Introduction to the foundations of causal discovery,” Int. J. Data Sci. Anal., vol. 3, no. 2, pp. 81–91, Mar. 2017.

[26] J. Peters, D. Janzing, and B. Schölkopf, Elements of Causal Inference. Cambridge, MA, USA: MIT Press, 2017.

[27] D. Malinsky and D. Danks, “Causal discovery algorithms: A practical guide,” Philos. Compass, vol. 13, no. 1, Jan. 2018, Art. no. e12470.

[28] C. Glymour, K. Zhang, and P. Spirtes, “Review of causal discovery methods based on graphical models,” Front. Genet., vol. 10, p. 524, Jun. 2019.

[29] K. Yu, J. Li, and L. Liu, “A review on algorithms for constraint-based causal discovery,” Nov. 2016, arXiv:1611.03977.

[30] J. Ramsey, M. Glymour, R. Sanchez-Romero, and C. Glymour, “A million variables and more: The fast greedy equivalence search algorithm for learning high-dimensional graphical causal models, with an application to functional magnetic resonance images,” Int. J. Data Sci. Anal., vol. 3, no. 2, pp. 121–129, Mar. 2017.

[31] Y. Song, F. Mao, and Q. Liu, “Human comfort in indoor environment: A review on assessment criteria, data collection and data analysis methods,” IEEE Access, vol. 7, pp. 119774–119786, 2019.

[32] N. Ma, D. Aiviv, H. Guo, and W. W. Graham, “Measuring the right factors: A review of variables and models for thermal comfort and indoor air quality,” Renew. Sustain. Energy Rev., vol. 135, Jan. 2021, Art. no. 110436.

[33] B. Sahoh, M. Kliangkhlao, and N. Kittiphatthananabowon, “Design and development of Internet of Things-driven fault detection of indoor thermal comfort: HVAC system problems case study,” Sensors, vol. 22, no. 5, p. 1925, Mar. 2022.

[34] J. L. Nguyen, J. Schwartz, and D. W. Dockery, “The relationship between indoor and outdoor temperature, apparent temperature, relative humidity, and absolute humidity,” Indoor Air, vol. 24, no. 1, pp. 103–112, Feb. 2014.

[35] J. F. Yin, Y. F. Zheng, R. J. Wu, J. G. Tan, D. X. Ye, and W. Wang, “An analysis of influential factors on outdoor thermal comfort in summer,” Int. J. Biometeorol., vol. 56, no. 5, pp. 941–948, Sep. 2012.

[36] D. Barber, Bayesian Reasoning and Machine Learning. 1st ed., Cambridge, U.K.: Cambridge Univ. Press, 2011.

[37] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, Bayesian Data Analysis, Boca Raton, FL, USA: Chapman Hall/CRC, 1995.

[38] X. Shen et al., “Challenges and opportunities with causal discovery algorithms: Application to Alzheimer’s pathophysiology,” Sci. Rep., vol. 10, no. 1, pp. 1–12, 2020.
Charlee Kaewrat received the B.S. degree in multimedia technology and animation, and the M.Sc. and Ph.D. degrees in management of information technology from Walailak University (WU), Nakhon Si Thammarat, Thailand, in 2011, 2016, and 2019, respectively.

He is currently a Lecturer with the School of Informatics, WU, and a member of the Informatics Innovation Center of Excellence. His current research emphasizes augmented reality, virtual reality, and human interaction.

Kirttayoth Yeranee received the B.Eng. degree in mechanical engineering from the Prince of Songkla University, Songkhla, Thailand, in 2014, and the M.Eng. degree in mechanical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2019, where he is currently pursuing the Ph.D. degree in mechanical engineering.

His research interests cover various topics, including heat transfer enhancement, computation fluid dynamic, and thermal performance enhancement.

Nichnan Kittiphattanabawon received the B.S. degree in computer science from Rangsit University, Pathum Thani, Thailand, in 1993, the M.Sc. degree in computer science from the Prince of Songkla University, Songkhla, Thailand, in 1999, and the Ph.D. degree in technology from Sirindhorn International Institute of Technology, Thammasat University, Bangkok, Thailand, in 2012.

She is currently an Assistant Professor with the Information Technology Department, School of Informatics, Walailak University, Tha Sala, Thailand. Her current research interests include text mining, data mining, document relation discovery, and association analytics.

Malika Kliangkhlao received the B.S. degree in information and communication technology from Silpakorn University, Bangkok, Thailand, in 2010, and the M.Sc. degree in information technology from Thaksin University, Phatthalung, Thailand, in 2014. She is currently pursuing the Ph.D. degree in computer engineering with the Prince of Songkla University, Songkhla, Thailand.

Her research interests cover various topics, including causal inference, explainable artificial intelligence, and modern supply chain management.